ABSTRACT

TZENG, SZ-TING. Understanding the Interplay of Social Signals and Values in Norm Emergence. (Under the direction of Munindar P. Singh.)

Advancements in technology have seamlessly integrated Artificial Intelligence (AI) into our daily lives. Software engages in dynamic interactions with its environment, other software, and human beings, fostering a symbiotic relationship. The interconnectedness gives rise to a multiagent system (MAS) where humans and AI work together in synergy to attain shared objectives. Given the involvement of humans, AI systems must be able to reason over human behaviors, which are determined by a combination of internal attitudes and external factors. Incorporating human values and considering multiagent dynamics in decision-making would lead to a substantial improvement in the reliability and realism of AI systems. Besides aligning decisions with values, humans have a fundamental need to comprehend and place trust in the output of AI.

Another concern that arises from the growing size and dynamics of the MAS is adaptability. Social norms define acceptable group conduct and governing agent behaviors in MAS. Norms can arise through top-down imposition or bottom-up emergence. In both approaches, norms and the environment are subject to change over time. The capacity for adaptation in AI systems becomes crucial to minimize human intervention and effort in maintaining MAS.

In this dissertation, we aim to incorporate adaptability and explainability in AI systems by integrating normative MAS with human factors. Initially, we concentrate on elements that help to regulate human behaviors. Subsequently, we explore human factors associated with human needs. This research includes four components: enforcing social norm with emotions, normative information from social signals, social value orientation, and decision and rationale with values. First, we introduce *Noe*, a framework that models the emotional responses of agents to the outcomes of interactions. Emotions, which are responses to internal or external events, can act as a positive or negative reinforcement mechanism for specific behaviors. Second, we introduce *Ness*, a framework that incorporates normative information from social signals to support norm emergence. In addition to sanctions, normative information from soft signals like hints and messages helps to regulate behaviors. Third, we present our *Fleur* framework, which incorporates the social value orientation concept. Social value orientation defines individuals' preferences over resource allocations between themselves and others. Aligning with social values enables AI to make ethical decisions and be responsible for human needs. Lastly, we describe *Exanna*, a framework that makes decisions and reveals information in rationales based

on agents' values. Value-aligned explanations ensure the AI system's decisions are consistent with human values and societal expectations.

© Copyright 2023 by Sz-Ting Tzeng

All Rights Reserved

Understanding the Interplay of Social Signals and Values in Norm Emergence

by
Sz-Ting Tzeng

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Computer Science

Raleigh, North Carolina

2023

APPROVED BY:

Arnav Jhala	Min Chi
William Rand	Nirav Ajmeri External Member

Munindar P. Singh
Chair of Advisory Committee

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to all those who have supported me throughout this remarkable journey of completing this dissertation.

First and foremost, I am profoundly thankful to my advisor Dr. Munindar Singh for his unwavering guidance, invaluable insights, and continuous encouragement. His expertise and dedication have been instrumental in shaping the direction and quality of this research. I genuinely appreciate my unofficial co-advisor, Dr. Nirav Ajmeri, for the continuous support and guidance he has provided.

I would also like to extend my sincere appreciation to my advising committee, Dr. Arnav Jhala, Dr. Min Chi, and Dr. William Rand, for their expert feedback and constructive critiques that significantly improved the quality of this work.

I am grateful to my friends and colleagues who provided a network of emotional support and academic inspiration during this journey.

My sincere thanks go to my family for their constant support and confidence in my abilities. Their love and support have been my driving force, and I am forever grateful.

Lastly, I thank the NSF (grant IIS-1908374 and grant IIS-2116751) for partial support for this research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER

1

INTRODUCTION

Advancements in technology have seamlessly integrated Artificial Intelligence (AI) into our daily lives. For example, Netflix's recommendation system suggests videos based on users' preferences; virtual assistants on smart devices process and execute user requests in natural language. In contrast to the past, software is now unrestricted by confined and isolated environments. With cutting-edge technology, software interacts dynamically with its surroundings, other software, and human beings (Kafalı et al. 2016). This symbiotic interaction leads to the establishment of a multiagent system (MAS), where a synergistic relationship emerges between humans and AI. Given the involvement of humans, it becomes crucial to incorporate human factors while constructing contemporary AI systems. Specifically, AI systems must be able to reason over human behaviors, which are determined by both internal attitudes and external factors. AI would become more reliable and realistic by incorporating human values and considering the multiagent dynamics in decision-making.

In Attribution theory, internal attributions explain human behavior with a focus on the characteristics of a person (Gerace 2020). e.g., their personalities, abilities, and physical characteristics. On the contrary, external attributions stress environmental or situational factors. e.g., social influences and task difficulty. In the theory of basic human values, values characterize individuals and societies (Schwartz 2012). Values explain behaviors and attitudes on a motivational basis. Specifically, human values define an individual's intrinsic motivation and dominate how this individual thinks and evaluates everything. We categorize values into two distinct categories: social values and individual values. Social values refer to the values of a society, while individual values delineate the values that characterize an individual. When a MAS becomes more interconnected, the complexity of interactions increases drastically, and it becomes hard to model all the possibilities. Basing on human values provides a solution to handle unexpected situations while aligning with human needs.

In addition to aligning decisions with values, humans have a fundamental need to comprehend and place trust in the output of AI. In other words, AI should be able to provide rationales for their decisions to be trusted by humans. However, explaining without considering individual differences may lead to information overload or privacy leaks among stakeholders.

Another concern that arises from the growing size and dynamics of the MAS is adaptability. Social norms play a significant role in MAS, defining acceptable group conduct and governing agent behavior (Savarimuthu and Cranefield 2011; Hollander and Wu 2011). A MAS incorporating norms that govern individual agents' behavior becomes a normative MAS. These norms elicit sanctions as responses to norm satisfaction or violation, e.g., penalties or rewards. Norms can arise through top-down imposition or bottom-up emergence (Morris-Martin et al. 2019). Top-down norms, e.g., laws and regulations, are costly and dictated by a central authority. Conversely, norms can also emerge from agent interactions in a bottom-up manner. In both approaches, norms and the environment are subject to change over time. The capacity for adaptation in AI systems becomes crucial to minimize human intervention and effort in maintaining MAS.

Sanctions, one form of social signals, coordinate and regulate agent behaviors in MAS. As humans evolved, social signals have emerged in the form of verbal messages or subtle hints, transmitting normative information.

While humans' decision-making includes internal and external attitudes, the other critical human factor in decision-making is emotions. Emotions, which are responses to internal or external events, can significantly impact decision-making and offer additional information in communication. Herbert Simon, a Nobel laureate, emphasized that general thinking and problem-solving must incorporate the influence of emotions (Simon 1967). Even more, emotions could be part of the norms themselves. Integrating both norms and emotions is essential for building explainable and trustworthy AI.

We aim to incorporate adaptability and explainability in AI systems by integrating normative multiagent systems (MAS) with human factors. Initially, we concentrate on elements that help to regulate human behaviors. Subsequently, we explore human factors associated with human needs. In this dissertation, we study methods to empower a normative MAS with the capability to adapt to dynamic environments and reason over human values.

1.1 Motivations

As AI systems increasingly involve humans in the decision-making process, there is a growing demand to consider human factors during their development. Varela et al. (2017) presents a contemplative aspect of human experience, among which we focus on social, cognitive, and emotional factors in this research.

1.1.1 Enforcing Social Norm with Emotions

In multiagent systems, norms and sanctions are often used to regulate agent behaviors while maintaining their autonomy. However, sanctions in the real world are more subtle instead of harsh punishment. For instance, the sanctions could be trust update or emotional expression and might change one's behavior (Nardin et al. 2016; Bourgais et al. 2019). At the basic level of Emotions' Social Functions, emotions help individuals understand others' preferences, beliefs, and intentions and coordinate social interactions (Keltner and Haidt 1999).

Consider a pandemic scenario. During a pandemic, many stores limit the number of customers in stores at once to protect their customers. A side effect of this practice is the long queue outside the stores. While there is a social norm that people should line up to enter the stores, some can still jump the queue to get services in advance. Suppose those who violate the norms would feel guilty (self-directed emotion) and receive negative emotions from others (other-directed emotion). These felt emotions will enforce the norm in stores.

The above scenario demonstrates the necessity of incorporating emotions when studying norm enforcement.

1.1.2 Modeling Normative Information via Social Signals

Social signals, as reactions to norm satisfaction or norm violation, provide natural drivers for norm emergence. When humans are evolved, social signals can be realized in three main ways: *sanction*, *tell*, and *hint*. Hints or emotions, as forms of non-verbal communication, are usually not considered in normative MAS. Normative information conveyed through a social signal also helps regulate MAS behaviors. In addition, hints may enable the inference of unobservable mental states (Wu et al. 2018; Wu and Schulz 2020).

With tell, agents communicate direct normative messages of approval or disapproval. An example of tell is verbal warning. An agent states clearly or indicates something may happen if someone does something. While messages provide clear normative information, hints also give subtle normative information for behavior. Upon receiving negative emotions after some actions, we can infer that our behaviors do not fit into others' expectations.

Consider the following example.

David notices Becka's suspicious symptoms and expresses some coldness near her. Upon perceiving David's negative attitude, Becka infers it is because she was sniffling near him in apparent violation of safety guidelines. Becka feels guilty about wandering out while being symptomatic. In addition, third parties who observe Becka's behavior and the actions of others may alter their behavior without directly having to be told.

The study of messages and hints as drivers of subtle social learning remains insufficient. Specifically, soft signals like hints have not been studied as drivers of norm emergence.

1.1.3 Social Value Orientation

While social norms regulate human behaviors, humans evaluate social norms based on human values and decide whether to comply or not. Social value orientation (SVO) is a psychological concept that describes individual differences in how people place value preferences along the dimensions of self and other. Consider there is a rare case scenario. During a pandemic, the authorities announce a mask-wearing regulation and claim that regulation would help avoid infecting others or being infected. Although Felix tests positive on the pandemic and prefers not to wear a mask, he also cares about others' health. If he stays in a room with another healthy person, Elliot, Felix will put the mask on.

While values may differ among individuals, a reliable AI system must take into account the values of its stakeholders to ensure making right decisions.

1.1.4 Decision and Rationale with Values

Two key aspects are essential for AI systems to earn human trust and be interpretable. Firstly, the systems must align their decisions with human values at the micro level (Liscio et al. 2023), focusing on the individual agent behaviors. While the macro level of values concerns the governance of MAS, the micro level ensures that decisions reflect individual human values. Secondly, agents should be capable of providing rationales for their decisions, enabling transparency and understanding behind their choices (Winikoff et al. 2021; Ayci et al. 2023). Rationales serve as vital information for making decisions and can aid in resolving social conflicts. However, determining the appropriate extent of information an agent should provide raises crucial questions. One one side, overly detailed rationales might become convoluted and fail to persuade, resulting in information overload. One the other side, disclosing private information could leads to potential privacy breaches.

When humans are involved in the MAS, it becomes crucial for AI systems to make decisions that extend beyond mere physical gain and, instead, harmonize with human values. As values play a significant role in guiding motivations and steering decisions, rationales aligned with values best justify one's behaviors. In addition, values reflect different concerns in decision-making and conflict resolution among agents. The above issues demonstrate the necessity of building rationales based on agents' value.

1.2 Research Objective and Questions

Based on the aforementioned challenges, the research objective that we aim to achieve is to design a framework that incorporates human factors and operates in dynamic environments, ensuring the trustworthiness of AI systems.

In order to achieve our research objective, we seek to address the following questions.

$\mathbf{RQ_{emotion}}$. How does modeling the emotional responses of agents to the outcomes of interactions affect the norm emergence?

$\mathbf{RQ_{signal}}$. How does considering hints and normative messages in addition to sanctions influence norm emergence?

$\mathbf{RQ_{SVO}}$. How do the preferences for others' rewards influence norm compliance?

$\mathbf{RQ_{rationale}}$. Do value-aligned rationales enrich the social experiences of agents?

1.3 Contributions

1.3.1 *Noe*: Enforcing Social Norm with Emotions

To address $\mathrm{RQ_{emotion}}$ in Section 1.1.1, we propose a framework *Noe* that integrates emotions in the normative reasoning process. Both norm satisfaction and violation elicit additional emotions, and these subsequent emotions impact the enforcement of norms.

1.3.2 *Ness*: Modeling Normative Information via Social Signals

To address the problems in Section 1.1.2, We present an agent framework that integrates social signals, encompassing sanctions, messages, and hints, to address our first research question RQ_{signal}. Our proposed framework *Ness* regards normative information from messages or hints as potential rewards that can shape behaviors. Including those signals enables indirect social learning, which resembles human behaviors in the real world.

1.3.3 *Fleur*: Social Value Orientation

Little research on norm emergence has incorporated social preferences, which shape the behavior of individuals when others are involved. To tackle the challenges in Section 1.1.3, we develop *Fleur*, a framework for agents that considers social value orientation and social norms while making decisions.

1.3.4 *Exanna*: Decision and Rationale with Values

For the challenges in Section 1.1.4, we propose *Exanna*, a framework that incorporates values in decision-making, rationale generation, and reasoning over rationale. While some research emphasizes the interpretability of agent decisions for humans, *Exanna* agents provide rationales to both agents and humans.

1.4 Organization

The dissertation is organized as follows.

Chapter 2 describes *Noe* and some related work. *Noe* shows how modeling the emotional responses of agents to the outcomes of interactions affect norm emergence and social welfare.

Chapter 3 introduces *Ness*, a framework that models normative information from social signals to support norm emergence. This chapter exhibits how agents with soft signals effectively avoid undesirable consequences, which are negative sanctions and deviation from goals, and yield higher satisfaction for themselves than baseline agents despite requiring an equivalent amount of social signals.

Chapter 4 presents our *Fleur* framework to address RQ_{SVO}. *Fleur* incorporates the social value orientation, which provides agents with different preferences over resource allocations between themselves and others. This chapter shows how social value orientation enables better social experience and robust norm emergence.

Chapter 5 describes *Exanna*, which addresses $\text{RQ}_{\text{rationale}}$ in simulated pandemic environments. This chapter demonstrates how value-aligned rationales enrich agents' social experiences.

Chapter 6 concludes this research and proposes possible future work.

CHAPTER

2

ENFORCING SOCIAL NORM WITH EMOTIONS

2.1 Introduction

Humans, in daily life, face many choices at many moments, and each selection brings positive and negative payoffs. In psychology, decision-making (Simon 1960) is a cognitive process that selects a belief or a series of actions based on values, preferences, and beliefs to achieve specific goals. Emotions, the responses to internal or external events or objects, can involve the decision-making process and provide extra information in communication (Keltner and Haidt 1999; Schwarz 2000). Social norms describe societal principles between agents in a multiagent system. While social norms regulate behaviors in society (Singh 2013; Savarimuthu and Cranefield 2011; Kafalı et al. 2020), humans and agents have the capacity to deviate from norms in certain contexts. For instance, people shake hands normally but deviate from this social norm during a pandemic. Chopra and Singh (2016) describe how social protocols rely on a foundation of norms though they do not discuss how the appropriate norms emerge.

An agent that models the emotions of its users and other humans can potentially behave in a more realistic and trustworthy manner. The decision-making process for humans or agents involves evaluating possible consequences of available actions and choosing the action that maximizes the expected utility (Edwards 1954). Herbert Simon, one of the founders of AI, emphasized that general thinking and problem-solving must incorporate the influence of emotions (Simon 1967). Without considering emotions or other affective characteristics, such as personality or mood, some compliance seems irrational (Argente et al. 2022). Humans' compliance shows hints on rational planning over their objectives (Keltner and Haidt 1999). Including emotion or personality in normative reasoning makes these compliance behaviors explainable. Norms either are defined in a top-down manner or emerge in a bottom-up manner (Savarimuthu and Cranefield 2011; Morris-Martin et al. 2019). Works on norms include norm emergence based on the prior outcome of norms, automated run-time revision of sanctions (Dell'Anna et al. 2020), or considering various aspects during reasoning (Ajmeri et al. 2020, 2018). However, sanctions in the real world are often subtle instead of harsh punishments. For instance, sanctions could be trust updates or emotional expression and might change one's behavior (Nardin et al. 2016; Bourgais et al. 2019). Kalia et al. (2019) considered norm outcome with respect to emotions and trust and goals. Modeling and reasoning about emotions and other affective characteristics in an agent then become important in decision making and would help the agent enforce and internalize norms.

Accordingly, we propose *Noe*, an agent architecture that integrates decision-making with normative reasoning and emotions. We investigate the following research question.

RQ$_{\textbf{emotion}}$. How does modeling the emotional responses of agents to the outcomes of interactions affect norm emergence and social welfare in an agent society?

To address RQ$_{\text{emotion}}$, we refine the abstract normative emotional agent architecture (Argente et al. 2022) and investigate the interplay of norms and emotions. We propose a framework *Noe* based on BDI architecture (Rao and Georgeff 1991), norm life-cycle (Savarimuthu and Cranefield 2011; Frantz and Pigozzi 2018; Argente et al. 2022), and emotion life-cycle (Alfonso Espinosa 2017, pp. 62–64) (Marsella and Gratch 2009). To evaluate *Noe*, we design a simulation experiment with various agent societies. We investigate how norms emerge and how emotions in normative agents influence social welfare.

To make the problem tractable, we apply one social norm in our simulation and simplify the emotional expression to reduce the complexity. Specifically, our *Noe* agents process emotions by appraising norm outcomes. For the emotion model, we adopt the OCC model of emotions (Ortony et al. 1988) in which we consider both emotional valence and intensity and assume violation of norms yields negative emotions.

Organization. The rest of the paper is structured as follows. Section 2.2 discusses the relevant related works. Section 2.3 describes *Noe*, including the symbolic representation and the decision-making in *Noe*. Section 2.4 details the simulation experiments we conduct to evaluate *Noe* and describes the experimental results. Section 2.5 presents the conclusions and the future directions.

2.2 Related Works

Ortony et al. (1988) model emotions based on events, action, and objects. Marsella and Gratch (2009) proposed a computational model of emotion to model appraisal in perceptual, cognitive, and behavioral processes. Moerland et al. (2018) surveyed emotions in relation to reinforcement learning. Keltner and Haidt (1999) differentiate the functional approaches and research of emotions by four-level analysis: individual, dyadic, groups, and cultural. Briefly, emotions provide some information for agents or people to coordinate social interactions. We take inspiration from these works.

Savarimuthu and Cranefield (2011) proposed a life-cycles model for norms and discussed varied mechanisms of norm study. Broersen et al. (2001) introduced the so-called Beliefs-Obligations-Intentions-Desires (BOID) architecture on top of the Beliefs-Intentions-Desires (BDI) architecture (Rao and Georgeff 1991), which further include obligation and conflict resolution. de Lima et al. (2019) developed Gavel, an adaptive sanctioning enforcement framework, to choose appropriate sanctions based on different contexts. However, these works do not consider emotions in the decision-making process.

Argente et al. (2022) propose an abstract normative emotional agent architecture, which combines emotion model, normative model, and Belief-Desire-Intention (BDI) architecture. Argente et al. defined four types of relationships between emotions and norms: (1) emotion in the process of normative reasoning, (2) emotion generation with norm satisfaction or violation, (3) emotions as a way to enforce norms, (4) anticipation of emotions promotes internalization and compliance of social norms. Yet, Argente et al. do not validate the interplay between emotions and norms with their proposed architecture.

Bourgais et al. (2019) present an agent architecture that integrates cognition, emotions, emotion contagion, personality, norms, and social relations to simulate humans and ensure explainable behaviors. However, emotions are predefined and not generated via appraisal in this work.

von Scheve et al. (2006) consider emotion generation with norm satisfaction or violation. Specifically, an observer agent perceives the transgression of a norm of another, its strong negative emotions (e.g., contempt, disdain, detestation, or disgust) constitute negative sanctioning of the violator. The negative sanctioning then leads to negative emotions (e.g., shame, guilt, or embarrassment) in the violator. Besides, compliance with the social norms can stem from the fear of emotional-driven sanctions, which would lead to negative emotions in the violator. Such fear enforces social norms. Yet, emotions are not part of the decision-making process in this work.

2.3 *Noe*

We now describe the architecture, norm formal model, and decision-making.

2.3.1 Architecture

Noe integrates the BDI architecture (Rao and Georgeff 1991) with a normative model (Argente et al. 2022; Frantz and Pigozzi 2018; Savarimuthu and Cranefield 2011) and an emotional model (Alfonso Espinosa 2017; Marsella and Gratch 2009). A *Noe* agent assesses the environment, including other agents' expressed emotions, its cognitive mental states, and infer possible outcomes to make a decision. Figure 2.1 shows the three components of *Noe*.

The normative component of *Noe* includes the following processes:

- Identification: the agent recognize norms from its norm base based on its beliefs

- Instantiation: activate norms related to the agent

- Normative reasoning process: the reasoning process makes decisions based on the beliefs, current intention, self-directed emotions, other-directed emotions received from others, active norms, and how the norm satisfaction or violation influences the world and itself The *Noe* agents then update the intention based on the results of normative reasoning

- Norm fulfillment process: check if a norm has been fulfilled or violated based on the selected action. The compliance or violation of a norm will then trigger an elicit emotion event that will be appraised at the emotion component

The BDI component includes the following parts:

- Beliefs: form beliefs based on perceptions

- Desires: generate desires based on the beliefs

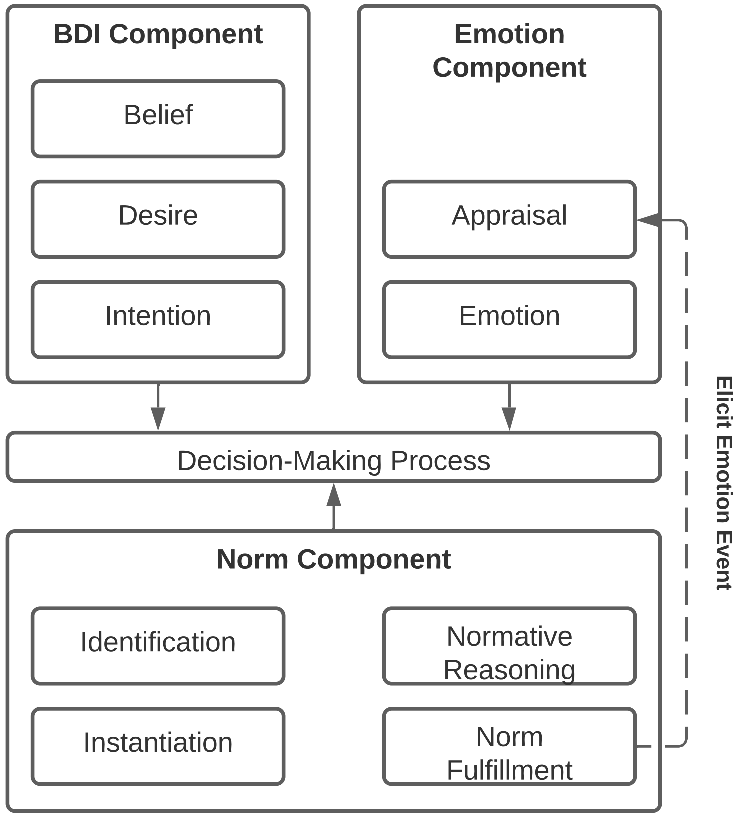


Figure 2.1: *Noe* architecture, representing and reasoning over beliefs, desires, intentions, emotions, and norms.

- Intention: the highest priority of desires to achieve based on the beliefs
- Action: select action based on the current intention, emotions, possible outcomes, and the evaluation of violating or complying with norms, if any

  The beliefs, desires, and intentions are mental states of *Noe* agents.

The emotional component includes the following processes:

- Appraisal: calculate the appraisal value based on the beliefs, desires, and norm satisfaction or norm violation. In this work, we consider only norm satisfaction or norm violation
- Emotion: generate emotion based on the appraisal values (Marsella and Gratch 2009)

Figure 2.2 illustrates the interactions between agents in our simulation scenario.

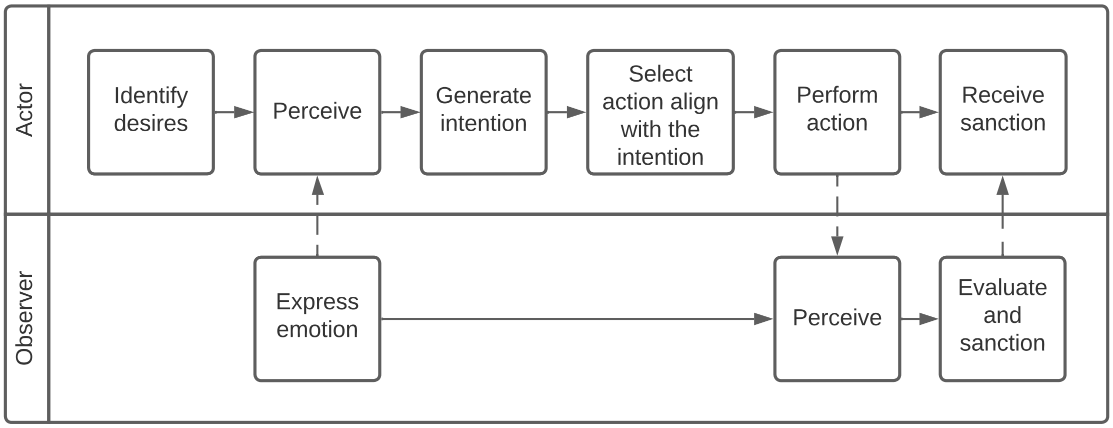


Figure 2.2: The interaction between *Noe* agents.

2.3.2 Norm Formal Model

Social norms describe the interactions between agents in a multiagent system. We adopt Singh's (2013) representation, where a social norm is formalized as $Norm(\text{subject}, \text{object}, \text{antecedent}, \text{consequent})$. In this representation, the subject and object represent agents, and the antecedent and consequent define conditions under which the norm is activated or satisfied, respectively. This representation describes a norm activated by the subject towards the object when the antecedent holds, and the consequent indicates if the norm was satisfied or violated.

Following Singh (2013), we consider three types of norms in *Noe*.

- Commitment (C): the subject commits to the object to bring out the consequence if the antecedent holds. Consider Alice and Bob are queuing up in a grocery store. Alice and Bob commit to keeping social distance during the pandemic, represented as $C(\text{Alice}, \text{Bob}, during = \text{pandemic}, \text{social distance})$.

- Prohibition (P): the object prohibits the subject from the consequence if the antecedent holds. Caleb, the grocery store manager, prohibits Bob from jumping the queue while lining up in that store, represented as $P(\text{Bob}, \text{Caleb}, when = \text{line up}; at = \text{grocery store}, \text{jump})$.

- Sanction (S): same as commitment or prohibition, yet the consequence would be the sanctions. Sanctions could be positive, negative, or neutral reactions to any norm satisfaction or violation (Nardin et al. 2016). If Bob breaks the queue, he receives negative sanctions from Alice, represented as $S(\text{Bob}, \text{Alice}, \text{jump}, \text{negative sanctions})$. Negative sanctions could be physical actions, e.g., scolding someone, or emotional expression, e.g., expressions of disdain, annoyance, or disgust.

To simulate the norm emergence and enforcement in human society, we include emotions into the decision-making process since, by nature, humans do not always act rationally in terms of utility theory. Here we formalize emotions with $E_i(target, intensity, decay)$ indicating agent a_i has emotion e toward the target with intensity and decay value. An example of the prohibition case would be, Bob would not jump the queue if Alice is angry, represented as $P(\text{Bob}, \text{Alice}, \text{Bob} \succ \text{Alice} \wedge E_{\text{Alice}} = \text{angry}, \text{jump})$.

We model the emotional response of agents with triggered emotions from norm satisfaction, or violation (Argente et al. 2022). Here we represent the elicited emotions with $Elem_{name}(A_{expect}, A_{real}, Em_1, Em_2) | Em_1, Em_2 \in E; A_{expect}, A_{real} \in A$ where A is a set of actions. E is a set of emotions, and Em_1 and Em_2 are the emotions triggered by norm satisfaction and violation accordingly. If the A_{expect} is equal to the A_{real}, a norm has been fulfilled, and Em_1 was elicited. $Ap(beliefs, desires, Elem)$ represents the appraisal function.

2.3.3 Decision-Making

Schwarz (2000) addresses the influence of moods and emotions at decision making and discusses the interplay of emotion, cognition, and decision making. Specifically, the aspects include pre-decision affect, post-decision affect, anticipated affect, and memories of past affect. In our model, we include the pre-decision affect into the decision-making process. With pre-decision affect, people recall information from memories that match their current affect (Schwarz 2000). For instance, people in a sad emotion or interacting with hostile people tend to overestimate adverse outcomes and events.

In our model, emotions serve as mental objects and an approach to sanctioning. We consider emotions as intrinsic rewards from agents' internal state in contrast to physical rewards from the environment. We adopt the OCC model of emotions (Ortony et al. 1988), in which we consider emotional valence and intensity. We formulate emotions with simple values where positive values indicate positive emotions and larger values indicate higher intensity. A mood is a general feeling and not a response to a specific event or stimulus compared to emotions. Therefore, we consider emotions but not mood. *Noe* agents' appraisal function considers norm satisfaction and violation only. The agents are aware of other agents' expressed emotions in the same place. In this work, we assume that agents express true and honest emotions and can correctly perceive the expressed emotions. In other words, felt emotions are equal to expressed emotions. Another assumption is that emotions are consistent with the notions of rational behavior.

Algorithm 1 displays the decision loop of our model. At the beginning of the simulation, all agents are initialized with certain desires, and during the run, an intention would be generated by prioritizing desires with the agent's beliefs. When choosing the next move with line 5 in Algorithm 1, the agent chooses the one with maximum utility from all available actions. Algorithm 2 details the action selection. The decision takes the agent's beliefs, current intention, and possible consequences into accounts. While norms are activated with the beliefs, the agent would further consider emotions and cost and possible consequences with norms at line 8 in Algorithm 2. For instance, if people violate some social norms, they may be isolated from society. Regarding the influence of emotions, people may overestimate the negative outcomes when they are in the negative emotion and tend to comply with the norms.

Algorithm 1: Decision loop of a *Noe* agent

1 Initialize one agent with its desires D;
2 **for** *t=1,T* **do**
3 Observe the environment (including the expressed emotions from others E_{around}) and form beliefs b_t;
4 Generate intention I based on b_t and D;
5 a_t = ActionSelection(b_t, I, D);
6 Execute action a_t;
7 Elicit self-directed emotions E_{self} from agent itself based on if action a_t fulfills a norm;
8 Self-sanction with E_{self};
9 Observe the environment (including the performed actions a_{t_other} of other agents) and form beliefs b_{t+1};
10 Elicit other-directed emotions E_{other} for observer agents based on if action a_{t_other} fulfills a norm;
11 Sanction others with E_{other};
12 **end**

2.4 Evaluation

We evaluate *Noe* via a line-up environment where agents form queues to receive service. We detail the environment in Section 2.4.1.

Algorithm 2: Action selection

Input: beliefs b_t, intention I, desires D
Output: Action a_t
Function `Action Selection`**:**
  $E_{around} \subset b_t$;
  for *each a in ACTIONS(b_t)* **do**
    Activate norms N with beliefs b_t and a;
    if $N = \varnothing$ **then**
      a_t = MAX$_a$(RESULT(b_t, intention, a))
    else
      a_t = MAX$_a$(RESULT(b_t, intention, a, N) $\times$ amplifier(E_{around}))
    end
  end
  return a_t
return

2.4.1 Line-Up Environment

Figure 2.3 shows the line-up environment. We build this line-up environment using Mesa (Masad and Kazil 2015), a Python-based framework for building, analyzing, and visualizing agent-based models.

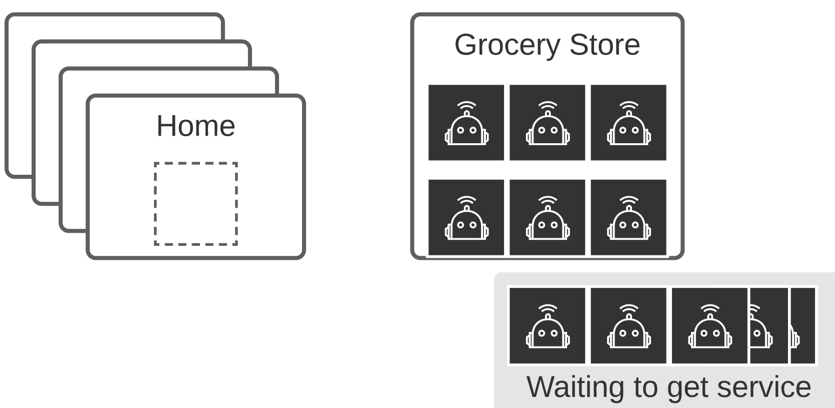


Figure 2.3: Simulation scenario. Agents move between their homes and the grocery store. The store has a capacity limit of eight customers at one time. As a result, other agents must line up outside the store to get service.

The line-up environment includes two shared locations—home and grocery stores. The agents move between home and grocery stores to get food. We consider one social norm in the line-up environment: agents are expected to line up to enter the grocery store. To simulate real human reactions to norm violations, we refer to a social psychology experiment (Milgram et al. 1986). In the line-up environment, we model defensive reactions of people in the queue as negative emotions toward those who jump the queue by barging in ahead of someone already in the queue. Conversely, people show positive emotions toward those who stay in the queue.

We initialize the agents with the following parameter values:

- Health (Integer value from 0–100): When the health value reaches zero, the agent is marked as *deceased* and unable to act. The health value decreases by 1 unit at each step.
- Deceased (Boolean: True or False): set as True when an agent runs out of health.
- Emotion (Integer value): simplified with numerical values where positive values indicate positive emotions and negative indicate negative emotions. The emotions come along with a duration. Default at 0.
- Number of food packets owned (Integer value from 0–15): once obtained food from the stores, agents would be able to restore its health value via consuming food anywhere.
- Food expiration day (Integer value from 0–15): once the agent gets food packets, we update the expiration day with 15. The expiration day decreases by 1 unit at each step. Food expires once the expiration day reaches 0. Default at 0.
- Beliefs: the perceived and processed information from the world, including other agents' expressed emotions.
- Desires: desired states, including *have food* and *wandering*.
- Intention: the highest priority of desires to achieve at a specific time. When the agent's health is lower than the threshold, 80% of the health, the agent sets its intention as *get food*; otherwise, the agent sets its intention as *wandering*.

When an agent runs low on stock, it has a higher probability of moving to a grocery store. The grocery store can provide food packets to eight agents in one time step. While waiting in line to get food, the agent could either stay in the line or jump ahead in the line to get food in less time. Jumping the line may increase other agents' delay in getting food packets. Those who witness the violation would then cast negative emotions, further interpreted as anger or disdain, triggered by that behavior. To simplify the simulation, we presume the anticipated affects (Schwarz 2000) with: (1) receiving negative emotions triggers negative self-directed emotions such as shame and guilt; (2) complying with norms leads to positive or neutral emotions; (3) violating norms leads to negative or neutral emotions. The intensity of emotions triggered each time is fixed, but the values of emotions can add up. Each triggered emotion lasts 2 steps. At each step, the duration and intensity of emotion decrease by 1 as decay. A simple assumption here is that people in a bad mood would trigger stronger emotions in response to a non-ideal state. Note that at the beginning of the simulation, we initialize the agent society with health in normal distribution to avoid all agents having the same intention at the same time.

2.4.2 Agent Types

To answer our research question and evaluate *Noe*, we define three agent societies as baselines. We describe the agents societies below:

Obedient society. Agents in an obedient society always follow norms.

Anarchy society. Agents in an anarchy society jump lines when they cannot get food.

Sanctioning society. Agents in the sanctioning society jump lines considering the previous experience of satisfying or violating a norm. Agents sanction positively or negatively based on norm satisfaction or violations directly and comply with enforced norms.

***Noe* society.** Agents in the *Noe* society jump lines considering the previous experiences of satisfying or violating a norm, current emotional state of the other agents, current self emotional state, and estimated outcome of satisfying or violating a norm. *Noe* agents who observe norm satisfaction or violations would appraise the norm outcomes and trigger emotions to sanction the actor agent.

Table 2.1 summarizes the characteristics of the agents in the four societies.

2.4.3 Hypotheses and Metrics

To address our research question $\mathrm{RQ}_{\mathrm{emotion}}$ on emotions and norm emergence, we propose three hypotheses:

$\mathbf{H_1}$ (Norm satisfaction): Norm satisfaction in *Noe* agent society is higher compared to the baseline agent societies.

Table 2.1: Characteristics of the various agent societies.

Agent Type	Violation allowed	Sanctioning	Emotions involved
Obedient society	✘	✘	✘
Anarchy society	✔	✘	✘
Sanctioning society	✔	✔	✘
Noe society	✔	✔	✔

H_2 (Social welfare): *Noe* agent society yields better social welfare compared to the baseline agent societies.

H_3 (Social experience): *Noe* agent society yields a better social experience compared to the baseline agent societies.

To evaluate H_1 on norm satisfaction, we compute one metric, M_1 (Cohesion): Percentage of norm satisfaction.

To evaluate H_2 on social welfare, we compute two metrics: (1) M_2 (Deceased): Cumulative number of agents deceased; (2) M_3 (Health): Average health of the agents. To evaluate H_3 on social experience, we compute one metric, M_4 (Waiting time): Average waiting time of agents in the queues.

To test the statistical significance of H_1, H_2, and H_3, we conduct the independent t-test and measure effect size with Glass' Δ for unrelated societies (Grissom and Kim 2012; Glass 1976). We adopt Cohen's (Cohen 1988, pp. 24–27) descriptors to interpret effect size where above 0.2, 0.5, 0.8 indicate small, medium, and large.

2.4.4 Experimental Setup

We run each simulation with 400 agents and queue size 80 for 3,000 steps. We choose a relatively small number of agents to reduce the simulation time while our results are stable for a more significant number of agents. The simulation stabilizes at about 1,500 steps, but we keep extended simulation steps to have more promising results. Table 2.2 lists the payoffs applied in our simulation.

We present the results with a moving average of 100 steps. We choose this size of running window to show the temporal behavior change in a small sequence of time. With a larger size, the running window may alleviate the behavior change. To minimize deviation from coincidence, we run each simulation with 10 iterations and compute the mean values.

Table 2.2: Payoff table.

Component	Type	Reward
Deceased	Extrinsic	–500
Norm compliance & positive emotion	Intrinsic	1
Norm violation & negative emotion	Intrinsic	–1

2.4.5 Experimental Results

In this section, we describe the simulation results comparing the three baselines and *Noe* agents. Table 2.3 summarizes the simulation results and the statistical analysis for our hypotheses.

According to Table 2.3, we see that *Noe* generate better cohesion and fewer deceased agents than baselines ($p < 0.01$; Glass' $\Delta > 0.8$). The null hypothesis corresponding to H_1 is rejected. Note that we do not consider the cohesion metric for the obedient agent society here since agents in the obedient society are always compliant. However, *Noe* also yields the worst social experience where the low waiting time is a desirable state ($p < 0.01$; Glass' $\Delta > 0.8$).

Table 2.3: Comparing *Noe* with baseline societies on various metrics and their statistical analysis with Glass' Δ and p-value. p is p-value from t-test.

		Obedient	Anarchy	Sanctioning	*Noe*
Cohesion	$\bar{X}$	–	0.22	0.88	**0.99**
	p	0.32	< 0.01	< 0.01	–
	Δ	0.19	102.43	13.67	–
Deceased	$\bar{X}$	55.30	81.60	169.30	**54.00**
	p	< 0.01	< 0.01	< 0.01	–
	Δ	0.65	3.10	15.53	–
Health	$\bar{X}$	79.27	79.50	86.26	79.00
	p	0.52	0.46	8.45	–
	Δ	0.18	0.21	3.34	–
Waiting time	$\bar{X}$	8.95	5.45	2.55	8.95
	p	0.98	< 0.01	< 0.01	–
	Δ	0.01	40.82	76.68	–

H_1 Norm Satisfaction

Figure 2.4 displays the cohesion, the percentage of norm satisfaction, in the baseline agent societies and the *Noe* agent society. We find that the percentage of norm satisfaction in the *Noe* agent society, average at 99% and p-value < 0.01, is constantly higher than the sanctioning agent society, average at 88% and p-value < 0.01 and Glass' $\Delta > 0.8$. The sanctioning agent society learns to comply with the norm as time goes by. The *Noe* agent society does sanction as well. Yet, considering emotions and the possible outcome makes *Noe* agent society enforce the norm faster than the sanctioning agent society. Specifically, *Noe* agent society enforces the norm at about 100 steps while sanctioning agent society at 1,500 steps.

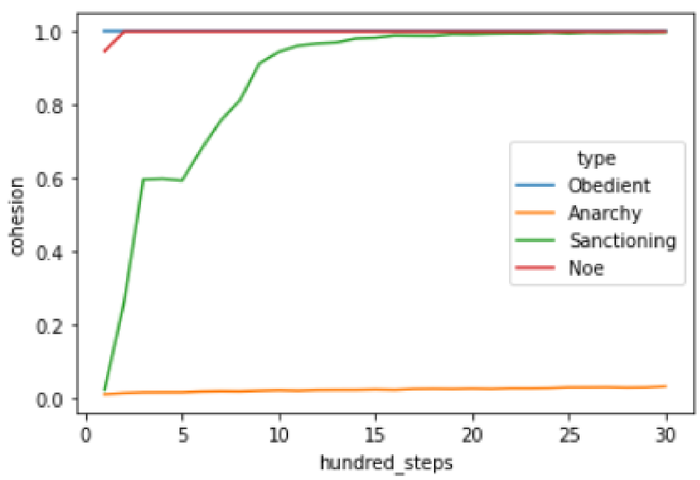


Figure 2.4: Simulation result: average cohesion. Comparing average cohesion (M_1) yielded by *Noe* and baseline agent societies.

H_2 Social Welfare

Figure 2.5 compares the average number of deceased in the obedient, anarchy, sanctioning, and *Noe* agent societies. Refer to Figure 2.4, sanctioning agent society learns the norm via positive and negative sanctioning from norm satisfaction and violation. However, the agents in that society do not consider the possible severe consequences and cause compliant agents to die in the queue. When the number of deceased reaches the threshold, the simulation stabilizes. Therefore, no more agent from the sanctioning agent society dies after the threshold. On the contrary, *Noe* agent society sanctions and considers possible outcomes of norm satisfaction and violation, therefore learning the norm and avoiding unacceptable consequences.

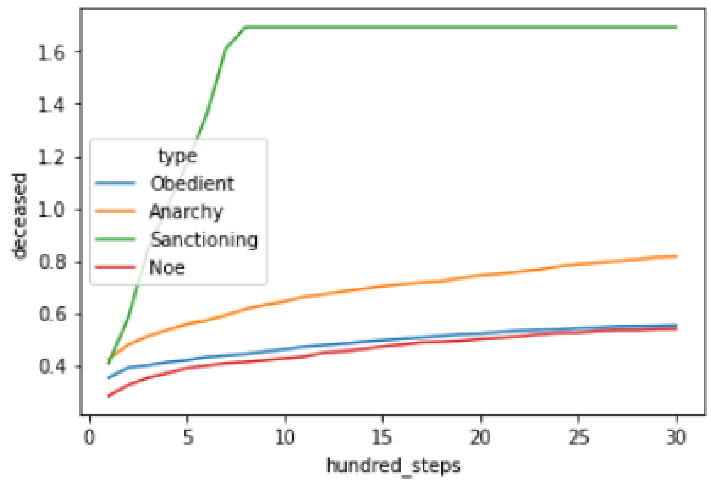


Figure 2.5: Simulation result: average number of deceased. Comparing average number of deceased (M_2) in *Noe* and baseline agent societies.

Figure 2.6 compares the average health of the agents in the obedient, anarchy, sanctioning, and *Noe* agent societies. The sanctioning agent society yields higher health State, with a mean at 86.26, but at the expense of more deaths. The rest of the agents then be able to remain in high health.

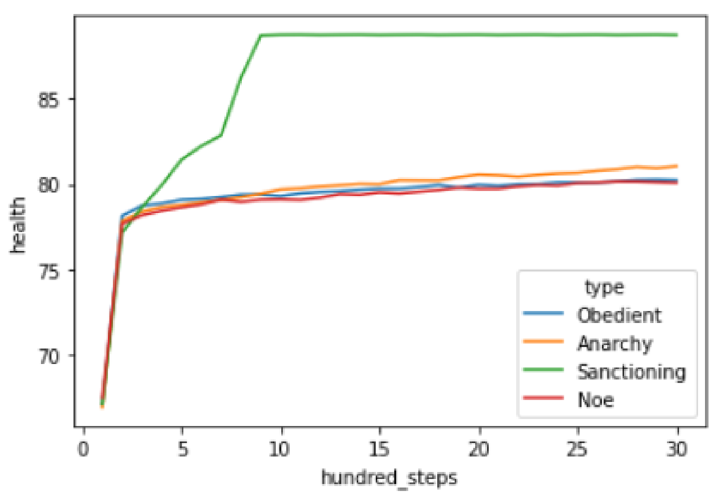


Figure 2.6: Simulation result: average health value. Comparing average health value (M_3) in *Noe* and baseline agent societies.

H_3 Social Experience

Figure 2.7 compares the average waiting time the agents spend in a queue at the grocery store in the obedient, anarchy, sanctioning, and *Noe* agent societies. The *Noe* agent society learns the norm fast and remains the same waiting time in the queue. However, some agents in the sanctioning agent society take advantage of those who learn norms faster than themselves. Therefore, many agents die during the learning process, and the simulation stabilizes. In Figure 2.7, the obedient agent society shares the same trend with *Noe* agent society since emotions enforce the line-up norm.

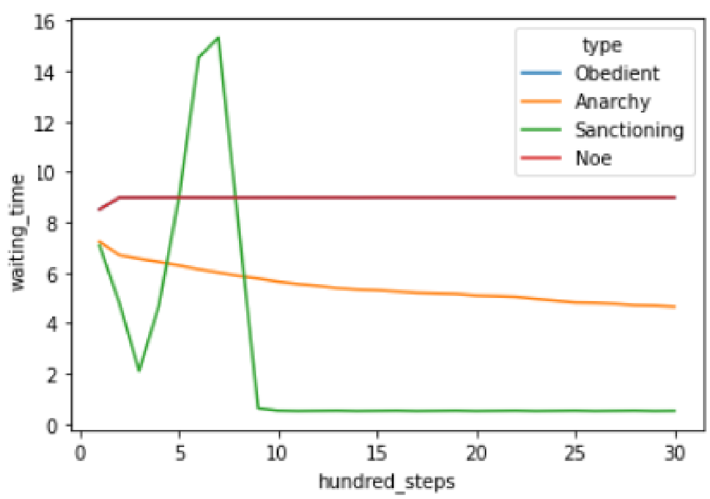


Figure 2.7: Simulation result: average waiting time of agents in queues. Comparing average waiting time (M_4) in *Noe* and baseline agent societies.

Combining the results for H_1 and H_2 and H_3, we note that while sanctioning enforces norms, a combination of sanctioning and emotions enforce norms better. Specifically, having emotions as amplifiers of outcomes yield higher norm satisfaction compared to our baselines. The results also indicate that, first, sanctioning agents that consider only norm violation or norm satisfaction may bring out worse social welfare compared to *Noe* that considers both norms and their consequences. Second, although *Noe* agents remain relatively high waiting time in the queues, the number of deceased is lower than the baselines. Note that the sudden drop of deceased number or increase of health value for sanctioning agents resulted from the stabilization of that society. Third, *Noe* agents stay in positive emotions during the simulation while sanctioning agents start from negative emotions and eventually achieve the expected behaviors.

2.5 Discussion and Conclusion

We present an agent architecture inspired by the norm life-cycle (Argente et al. 2022), BDI architecture (Rao and Georgeff 1991), and emotion life-cycle (Alfonso Espinosa 2017; Marsella and Gratch 2009) to investigate how emotions influence norm emergence and social welfare. We evaluate the proposed architecture via simulation experiments in an environment where agents queue up to receive service. Our simulations consider two characteristics of an agent society: sanctioning and emotions that participate in action selection and arise from evaluating selected action. The experiments show that incorporating emotions enables agents to cooperate better than those who do not.

In our agent architecture, we assume that agents can recognize others' emotions. However, we acknowledge that emotion recognition is a challenging task (Barrett et al. 2019). Whereas recent works in AI have focused on emotion recognition through facial expressions and emotion recognition using wearables, it is worth noting that there is no agreement in modeling emotions in the psychology community (Barrett et al. 2019; Marín-Morales et al. 2018; Marsella et al. 2010).

Murukannaiah et al. (2020) address many shortcomings of current approaches for AI ethics, including taking the value preferences of an agent's stakeholder and other agents' users, learning value preferences by observing the responses of other agents' users, and value-based negotiation. Incorporating these aspects in *Noe* is an interesting future direction.

As a future extension of current work, we intend to distinguish between emotions in *Noe* rather than representing them through emotional valence. This extension aims to enhance the depth of information available for understanding value preferences. We are also contemplating the incorporation of a variety of personalities in upcoming studies to yield diverse appraisal outcomes. In this work, *Noe* agents are assumed to express true and honest emotions. However, emotions can also serve as a tool to influence, persuade, or deceive others in an adversarial context. It would be crucial to identify and model these contradictions while humans are in the loop.

CHAPTER

3

MODELING NORMATIVE INFORMATION VIA SOCIAL SIGNALS

3.1 Introduction

Social norms characterize collective and acceptable group conduct and regulate agent behavior. Norms may be imposed top-down (as legal norms are) or emerge bottom-up (when agents learn acceptable behaviors from each other) (Savarimuthu and Cranefield 2011). Our interest is in the latter while accommodating the former. A norm emerges in a society when a substantial majority of its agents the same action in the same circumstance (Morris-Martin et al. 2019; Savarimuthu and Cranefield 2011).

We posit that the emergence of norms is driven by three kinds of social signaling by one agent to another in response to the first agent observing certain behaviors by the second agent in certain situations: (1) *Sanctions* or punishments or rewards (Nardin et al. 2016) for observed behaviors, (2) *Tell* or direct normative messages or explicit communications of approval or disapproval (Andrighetto et al. 2013) of observed behaviors, and (3) *Hint* or implicit signals conveying a positive or a negative attitude toward an observation.

Example 1 ***Sanction.*** *Becka is symptomatic with COVID-19. Alice meets Becka in a cafe and notices Becka's symptoms. Alice reports the violation of healthcare guidance, leading to Becka being required to undergo compulsory quarantine at designated facilities.*

Example 2 ***Tell.*** *Charlie notices Becka's suspicious symptoms and begins to worry about his safety. He tells Becka that roaming in public while symptomatic will be required to undergo compulsory quarantine.*

Example 3 ***Hint.*** *David notices Becka's symptoms and expresses some coldness near her. Becka interprets David's negative attitude as a reaction to her apparent violating safety guidelines by sniffing near him. Becka feels guilty for wandering out while being symptomatic and is anxious about the possibility of being reported.*

The above social signals convey normative information from which Becka learns her behavior was inappropriate. And, third parties who observe Becka's behavior and these signals may alter their behavior without directly having to be told.

Messages and hints drive subtle forms of social learning, as in human societies, but have not been adequately studied. Hints are soft signals that have not been studied as drivers of norm emergence. We investigate the following research question.

$\text{RQ}_{\text{signal}}$. How does considering hints and normative messages in addition to sanctions influence norm emergence?

To address $\text{RQ}_{\text{signal}}$, we define two expressions of normative information: explicit normative message (Andrighetto et al. 2013) and implicit hint as information.

Contributions. We propose *Ness* (for *Norm Emergence through Social Signals*), a framework that accommodates norms imposed top-down and enables norm emergence. *Ness* includes normative information from three types of social signals; that information facilitates social learning.

We evaluate *Ness* experimentally via a simulation of a pandemic scenario. We examine societies characterized by three distinct signal types: sanction, tell or direct messaging, and hint. Our results demonstrate that introducing normative information communicated via hints and direct messaging enables faster norm emergence, avoids undesirable consequences such as negative sanctions and deviation from goals, and yields higher satisfaction overall in a society. Especially, in societies with low vaccination rates, individuals learn that engaging in self-isolation is praiseworthy and short-term compromises prevent major penalties.

Organization. Section 3.2 introduces key concepts of *Ness* and describes how agents' decision-making work in *Ness*. Section 3.3 details the pandemic simulation we create to evaluate *Ness*.

Section 3.4 presents results from our simulation experiments. Section 3.5 discusses other relevant research. Section 3.6 concludes with a summary of our findings, limitations, threats to validity, and future directions.

3.2 *Ness*

A *Ness* agent selects actions considering their goals, environmental norms, and social signals (Argente et al. 2022; Frantz and Pigozzi 2018; Savarimuthu and Cranefield 2011; Singh 1994; Marsella and Gratch 2009).

A *Ness* agent learns from observations. Following Example 2, on receiving Charlie's message, Becka learns that she may be reported to local authorities if she does not self-isolate. In Example 3, Becka may have misread David's coldness as directed at her.

When making decisions, a *Ness* agent activates those norms related to itself based on its knowledge. The normative reasoning process enables *Ness* agents to reason over the possible outcome of norm compliance or violation. After executing a chosen action, the agent checks if a norm has been fulfilled or violated. The compliance and violation of norms then trigger social signals.

3.2.1 Key Concepts

We now introduce the key concepts in *Ness*.

Goal is a condition that an agent wants to achieve. The outcome of a goal has a binary value, achieved or not, after performing a selected action.

Norm defines the relationship between an agent on whom the norm is focused and an agent with respect to whom the norm arises. An agent can invest effort on its norm or has its freedom curtailed by another agent. A norm can either be satisfied or violated when the consequent holds or not, respectively.

Sanction refers to a positive, negative, or neutral reaction directed from one agent toward another. A sanction is typically in response to a norm satisfaction or norm violation.

Tell or normative message specifies the cause and the effect. The effect describes a potential reward or punishment. For example, Charlie's specification in Example 2 includes whether a norm is satisfied or violated and why.

Hint is an indirect clue that an agent expresses toward another agent to guide its behavior. A hint requires the receiver to infer the intended meaning. We model hints as subtle soft signals triggered by norm satisfaction or violation.

Reward Shaping refers to supplemental or "shaping" rewards (in addition to those from the environment) provided to agents to move toward a certain goal or to encourage selecting a certain action in a certain set of states (Marom and Rosman 2018). Here, we consider normative information from tell or hint as advice on potential soft sanctions with different levels of certainty. That is, signals—*tell* and *hint*—are inferred as positive or negative rewards to encourage or discourage taking specific actions.

3.2.2 Decision-Making

An agent's behaviors include acting to maximize its payoff, and giving social signals.

Action selection An agent selects an action that satisfies its goal and maximizes its actual and possible payoff. In the examples of Section 3.1, Becka decides whether to go to the cafe depending on her goals and her understanding of norms.

Social signal expression An agent observes other agents' behaviors and expresses social signals: sanctions, messages, or hints if the behaviors conflict with norms. In Example 1, Alice sanctions Becka based on a healthcare guidance of staying at home when symptomatic. In Example 2, Charlie sends Becka a direct message. In Example 3, David's coldness toward Becka shows his disapproval.

Reward Shaping A message or hint serves as a look-ahead advice on what will happen after a specific action. A shaping reward can be defined as $r' = r + F$ where r is the original reward function, and F is the shaping reward function. With messages or hints, F defines the difference of potential values. Here, Φ is a potential function that gives hints on states. κ defines the certainty of potential reward from the knowledge or information.

$$F(s,a,s',a') = \gamma\Phi(s',a')\kappa - \Phi(s,a) \tag{3.1}$$

3.3 Simulation

We evaluate *Ness* via a simulated pandemic scenario where how agents behave influences the spread of a pandemic. We built our pandemic environment using Mesa (Masad and Kazil 2015), a Python-based simulation framework. Our focus is not to model the realism of pandemic spreading but to investigate how social signals influence norms. Agents in the simulation use reinforcement learning to learn the relationship between objectives and normative behaviors.

3.3.1 Pandemic Scenario

In the simulated environment, an agent moves between four places: home (unique for each agent), park, cafe, and (vaccination) clinic. An agent has a goal to rest, hike, shop, be_vaccinated, and selects actions from {stay_home, visit_park, visit_cafe, visit_clinic}. Any two agents present at the same place may interact with equal probability at each step. Agents perceive each other's signals, and all expressed signals are genuine and honest.

At each step, an agent observes its environment and moves based on factors such as based on which an agent learns where to move include death, goal satisfaction, sanctions, messages, hints, and norm satisfaction or violation. After all agent move, they evaluate the behavior of, and accordingly signal, each other. An agent who witnesses another agent being signalled can learn from it.

3.3.2 Disease Model

Our disease model is simplified from the Susceptible-Exposed-Infected-Recovered (SEIR) model (Yang and Wang 2020; Annas et al. 2020) and captures the effectiveness of vaccines. As shown in Figure 3.1, each agent begins in a healthy state. Upon encountering an infected agent (not shown), it transitions to the asymptomatic phase of the disease, showing no symptoms despite being infected. As the symptoms progress, an agent becomes mildly symptomatic, then critically symptomatic, and in the worst-case, deceased. Vaccination offers protection by reducing the probability that agents can become infected and advance toward critical symptomatic or deceased. Home-based recovery is the primary treatment during the pandemic.

We base the probabilities of how COVID-19 evolves on Poletti et al. (2020). We set the infection probability to 80% and the effectiveness of vaccination at 50% to represent a more infectious variant and speed up the simulation. Apart from vaccination, we set the probability of the symptoms to progress as Figure 3.1. The intuition is that each infected person provides an opportunity for the symptoms to progress to the next phase or recover.

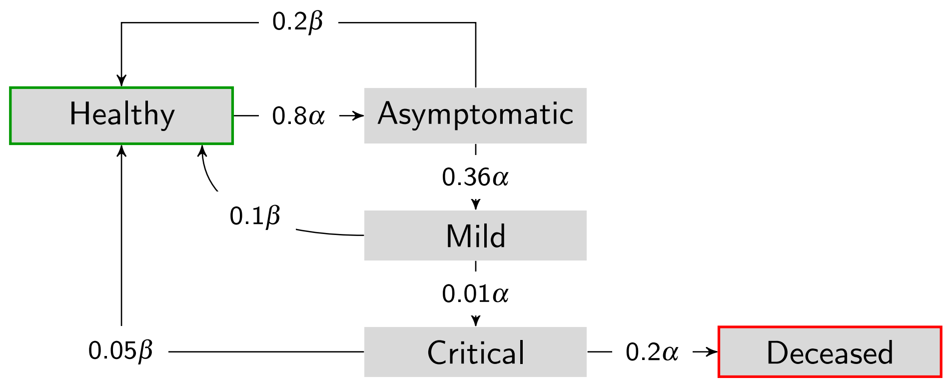


Figure 3.1: Disease model with state transition probabilities. The transition for healthy agents applies when coming in contact with those who are infected. α and β parameterize the transition probabilities for vaccination and home rest, respectively. In our study, $\alpha = 0.5$ means vaccinated, and $\alpha = 1.0$ means not vaccinated. And, $\beta = 2.0$ means at home and $\beta = 1.0$ means not at home. For example, the edge weight from Mild to Critical can be read as 0.01 for vaccinated agents and 0.005 for unvaccinated agents. The probability of remaining in the state is $1-$ the probability of evolving to the next state.

We assume home stay improves recovery from infection. We write "isolation" when home stay is voluntary and "quarantine" when it is forced. We set a 50% probability to sanction those who exhibit mild symptoms and an 80% probability to sanction those who exhibit critical symptoms but are not isolated. Table 3.1 shows partial observability of the health states of others (e.g., a healthy person who has watery eyes because of pepper may be perceived as sick (Mild)).

Table 3.1: Imperfect observation of another's health state.

Actual \ Belief	Healthy	Mild	Critical
Healthy	0.80	0.10	0.10
Asymptomatic	0.50	0.50	0.00
Mild	0.30	0.60	0.10
Critical	0.10	0.30	0.60

3.3.3 Social Norm

We initialize the environment with a social norm that healthy agents prohibit infected agents from staying in a public space. We frame the norm as a prohibition as below.

```
norm type   = {Prohibition},
subject     = {Infected_Agent},
object      = {Healthy_Agent},
antecedent  = {obs_health=[MILD, CRITICAL]},
consequent  = {loc=[PARK, CAFE, CLINIC]}
```

When the antecedent and consequent both hold, the prohibition is violated, and a sanction is given to the subject. The sanction is a numerical reward to the agent who violates the prohibition. When an agent receives normative information from others indicating or stating this prohibition, it considers the sanction a potential reward Φ. According to the signal type, the agent constructs the shaping reward.

3.3.4 Normative Information Communication

We formalize normative information via conditionals indicating that the stated consequent will be brought out when the presented antecedent holds. For example:

```
sender     = {Observer_Agent},
receiver   = {Actor_Agent},
info type  = {MESSAGE},
antecedent = {obs_health=CRITICAL,loc=CAFE},
consequent = {PUNISHMENT}
```

3.3.5 Types of Societies

We consider societies based on the social signal types: sanctioning, telling or direct messaging, and providing hints.

Baseline 1: PRIMITIVE society

Agents apply no social signals and act solely based on goal satisfaction (payoffs).

Baseline 2: PENALTY society

Agents obtain negative sanctions for violating a social norm, as in Example 1. Healthy agents may punish infected agents who enter a public space by directing them to quarantine. In the next few timesteps, a punished agent's position is changed to home regardless of its wishes.

Baseline 3: EMOTE society

This society is a variant of *Ness* but without shaping rewards. With some social norm in mind, agents who violate or satisfy the social norm receive a signal, such as expressed emotions, from others. These agents may experience guilt or pleasure based on norm violation or satisfaction. EMOTE is adapted from (Tzeng et al. 2021). Infected agents who wander in a public space may receive expressed emotions and feel bad about their norm violation and may be forced to quarantine at home by healthy agents.

Baseline 4: TELL society

This society is a variant of *Ness* but without the *hint* part, as in Example 2. Agents learn social norms and convey normative messages upon witnessing a norm violation. The normative message is adapted from (Andrighetto et al. 2013) and includes what sanctions an agent will receive if it violates a norm. Healthy agents interacting with infected agents in a public space convey the social norm of staying away from public spaces to the infected agents. Also, infected agents in a public space may be forced to quarantine by healthy agents.

Ness: Hint Society

A society with our proposed agents. *Ness* agents learn norms from social signals of sanction and hint. In the simulation, infected agents who wander in public spaces receive hints from healthy agents via which they infer the normative information about staying away from healthy agents. This information from the hint signal provides shaping rewards to agents to learn norms. *Ness* agents may also experience pleasure or guilt based on norm violation or satisfaction. In addition, as in PENALTY, EMOTE, and TELL, infected agents in a public space may be forced to quarantine at home by healthy agents.

3.3.6 Metrics

We compute six measures to evaluate *Ness*. M_{Healthy}, M_{Infected}, M_{Deceased}, $M_{\text{Infections}}$, $M_{\text{Vaccinated}}$, and M_{Goal} help identify the consequences of agents' behaviors or norm emergence. Moreover, these measures provide insights into why specific norms emerge. M_{Home} and $M_{\text{Quarantine}}$ yield the percentage of self-isolation behaviors.

$\mathbf{M_{Healthy}}$ Percentage of agents who are healthy.

$\mathbf{M_{Infected}}$ Percentage of agents who are infected.

$\mathbf{M_{Deceased}}$ Percentage of agents who are deceased.

$\mathbf{M_{Infections}}$ Average number of infections.

$\mathbf{M_{Vaccinated}}$ Percentage of agents who are vaccinated.

$\mathbf{M_{Home}}$ Percentage of infected agents who stay home.

$\mathbf{M_{Quarantine}}$ Number of agents forced to quarantine at home. This measure maps to the sanction signal type.

$\mathbf{M_{Goal}}$ The average goal satisfaction among agents.

A norm emerges when the proportion of agents following the same behavior exceeds a threshold. We consider 90% as the threshold (Delgado 2002).

3.3.7 Hypotheses

We evaluate three hypotheses. For each, we test statistical significance with respect to its null hypothesis via the independent t-test. We adopt Glass' Δ to measure the effect size since the standard deviations are different between societies (Glass 1976; Grissom and Kim 2012). We consider Cohen's (1988) descriptors to interpret the effect size. Specifically, an effect size less than 0.2 indicates that the difference is negligible; [0.2–0.5) indicates small; [0.5–0.8) indicates medium; and, 0.8 or above indicates a large effect.

$\mathbf{H_{Disease\ control}}$. Societies considering hints have better control over disease spread than the societies that do not consider hints. We compare societies with respect to $\mathrm{M_{Healthy}}$, $\mathrm{M_{Infected}}$, $\mathrm{M_{Deceased}}$, $\mathrm{M_{Infections}}$, and $\mathrm{M_{Vaccination}}$.

$\mathbf{H_{Isolation}}$. Societies considering hints yield improved isolation than other societies. We compare societies with respect to $\mathrm{M_{Home}}$, $\mathrm{M_{Quarantine}}$, and $\mathrm{M_{Infected}}$.

$\mathbf{H_{Goal}}$. Agents in *Ness* have higher goal satisfaction than other societies. We compare societies with respect to $\mathrm{M_{Goal}}$.

3.3.8 Experimental Setup

Table 3.2 lists the elements of reward function for an agent, including extrinsic rewards from the environment and intrinsic rewards from an agent's internal state. For data efficiency, we apply policy parameter sharing (Gupta et al. 2017) based on the assumption of the bystander. We consider 100 agents (30 are infected initially) with the simulated world lasting for 2,000 steps. Each society stabilizes within 1,500 steps. We train our agents for 100,000 steps, and report results averaged over 20 runs.

We consider normative information as shaping rewards that are part of intrinsic rewards. Specifically, we incorporate knowledge of being punished in the future from tell or hint into our simulation. Table 3.2 lists elements based on *Ness* agents appraises their states.

Table 3.2: Reward function. Sanctioning means quarantine. An agent's extrinsic rewards come from the environment and intrinsic rewards come from its current internal state. Norm satisfaction or violation is based on the action and the perceived health state of others instead of the actual health state.

Component	Type	Reward
Deceased	Extrinsic	–2
Sanctioning	Extrinsic	–1
Goal satisfaction	Intrinsic	+1
Goal violation	Intrinsic	–1
Norm satisfaction (self)	Intrinsic	+0.5
Norm violation (self)	Intrinsic	–0.5
Norm satisfaction (other)	Extrinsic	+0.5
Norm violation (other)	Extrinsic	–0.5

Information Balance

As messages and hints provide additional normative information to learn, we keep the signal distribution at the same level to balance the amount of information agents receive from a combination of signals.

Since enhancing the signals can improve learning, we adjust the signal distributions to balance the information across the societies. Table 3.3 lists the probability distribution over the various types of signal we apply for each society.

In Table 3.3, *Sanction* is the probability of an agent being compelled to quarantine. An agent considers a sanction as a punishment. *Tell* is the probability that an agent receives messages from its neighbors. An agent consider *tell* as potential rewards or punishments with 50% probability. *Emote* is the probability that an agent receives implicit signals such as expressed emotions conveying a positive or a negative attitudes as subtle sanctions. *Hint* is the probability that an agent receives positive or negative attitudes as implicit signals. The agent infers the signals as potential positive or negative rewards to encourage or discourage specific behaviors with 30% probability.

Table 3.3: Signal distributions. In each society, agents send a combination of the three social signals. To balance the amount of information, we adjust the distributions immediate (I) and potential (P) rewards as here. Here, w_i and w_p describe the weights associated with the rewards. EMOTE expresses other-directed attitudes towards others' behaviors as sanctions and has a self-directed attitude toward the self.

	Signals					w_{I}	w_{P}
	Sanction	Tell	Emote	Hint	None		
Signal type	I	P	I	I+P			
PRIMITIVE	0%	0%	0%	0%	100%	0	0
PENALTY	38%	0%	0%	0%	62%	1	0
TELL	20%	18%	0%	0%	62%	1	0.5
EMOTE *	20%	0%	18%	0%	62%	$1+0.5$	0
Ness: Hint	20%	0%	0%	18%	62%	$1+0.5$	0.3

Reinforcement Learning Parameters for Social Signals

Ness and baseline agents employ Q-Learning (Watkins and Dayan 1992) to learn norms. Q-Learning is a model-free reinforcement learning algorithm that learns from trial and error. The Q-Learning algorithm computes the action-state value $Q(s,a)$ (Q value), which indicates the expected and cumulative rewards for each state and action. By approximating the value of an action for a given state, the Q-Learning algorithm finds the optimal policy. The Q function computes Q values with the weighted average of the old value and the new information, and is given by:

$$Q(s,a) = Q(s,a) + \alpha \times (r_t + \gamma \max_{a'} Q(s',a') - Q(s,a)) \tag{3.2}$$

where $Q(s,a)$ is the expected value for performing action a in state s. Here, α is the learning rate and γ is the reward discount rate. s' refers to the next state, and a' refers to possible actions in s'.

Messages give precise causality between claimed behaviors and possible sanctions. On the contrary, hints provide subtle normative information for behaviors, which requires further inference. While hints and messages provide different levels of certainty of possible sanctions, we model normative information with approving and disapproving attitudes as shaping rewards. Specifically, we associate approving and disapproving attitudes with norm satisfaction and violation. The potential rewards are calculated from the signal based on the signal type. We set κ as 0.30 for hints and κ as 0.50 for messages, where κ is the certainty of possible sanctions from normative information. The supplementary material provides simulation hyperparameters for reproducibility.

3.4 Experimental Results

We now discuss the results for our research question $\text{RQ}_{\text{signal}}$. Table 3.4 summarizes the simulation results and the corresponding statistical analysis for $\text{RQ}_{\text{signal}}$. For each hypothesis, the metric reported is upon convergence.

3.4.1 $\mathbf{H}_{\textbf{Disease control}}$

To evaluate $\text{H}_{\text{Disease control}}$, we measure the proportion of healthy ($\text{M}_{\text{Healthy}}$), infectious ($\text{M}_{\text{Infected}}$), and deceased ($\text{M}_{\text{Deceased}}$) agents. We also track the average number of infections ($\text{M}_{\text{Infections}}$) and vaccination rate ($\text{M}_{\text{Vaccinated}}$). Infectious agents include those who are asymptomatic, mild symptomatic, and critical. Figure 3.2 reports these metrics. These simulation start from a 30% infection rate in each society. First, *Ness* has a lower fraction of infected agents (0.22) than PRIMITIVE (13.29), PENALTY (2.65), EMOTE (3.78), and TELL (2.96). The effect is large for PRIMITIVE and small for PENALTY and TELL but negligible for EMOTE.

Second, *Ness* has the more healthy agents (97.54) than PRIMITIVE (46.31), PENALTY (77.60), EMOTE (67.11), and TELL (76.27). The effect is large.

Third, *Ness* has a lower fraction of deceased agents (2.08) than PRIMITIVE (41.01), PENALTY (19.75), EMOTE (29.10), and TELL (20.78). *Ness* has a lower $\text{M}_{\text{Infections}}$(2.07) than PRIMITIVE (48.31), PENALTY (13.83), EMOTE (19.09), and TELL (15.16). The effect is large for each case.

Table 3.4: Comparing *Ness* with baseline societies on various metrics and their statistical analysis with Glass' Δ. All p-values are <0.001. The metrics reported are calculated upon convergence. The performance in disease control is based on attitudes expressed and information shared as in the order of *Ness*, EMOTE, TELL, PENALTY, and PRIMITIVE. However, the results of vaccination are in reverse order. *Ness* have higher isolation and quarantine rate than TELL, EMOTE, and PENALTY. Goal satisfaction ranked in this order *Ness*, TELL, PENALTY, EMOTE, and PRIMITIVE.

		PRIM.	PEN.	EMOTE	TELL	*Ness*
$H_{\text{Disease control}}$	M_{Infected}	13.29	2.65	3.78	2.96	0.22
	Δ	0.96	0.25	0.30	0.26	–
	M_{Healthy}	46.31	77.60	67.11	76.27	97.54
	Δ	17.54	3.22	4.79	3.23	–
	M_{Deceased}	41.01	19.75	29.10	20.78	2.08
	Δ	3.25	5.74	5.27	5.48	–
	$M_{\text{Infections}}$	48.31	13.83	19.09	15.16	2.07
	Δ	2.59	6.21	5.56	6.00	–
	$M_{\text{Vaccinated}}$	82.41	36.72	32.69	35.33	93.57
	Δ	1.03	16.21	15.86	14.78	–
$H_{\text{Isolation}}$	M_{Home}	0.61	0.96	0.95	0.95	0.99
	Δ	1.76	0.29	0.38	0.35	–
	$M_{\text{Quarantine}}$	–	0.03	0.02	0.02	0.00
	Δ	–	0.26	0.28	0.26	–
H_{Goal}	M_{Goal}	0.19	0.26	0.23	0.26	0.31
	Δ	3.04	3.01	3.67	3.04	–

With regard to $M_{\text{Vaccinated}}$, *Ness* has a higher vaccination rate (93.57) than PRIMITIVE (82.41), PENALTY (36.72), EMOTE (32.69), and TELL (35.33). The effect is large.

With $M_{\text{Vaccinated}}$, we observe that a vaccination norm emerges with a majority above 90% in *Ness*. Specifically, even without a top-down imposed shared expectation on vaccination, most agents in *Ness* learn that vaccination maximizes their payoff. The emerged vaccination norm can be articulated as below.

```
norm type   = {Committment},
subject     = {Infected_Agent},
object      = {Healthy_Agent},
antecedent  = {obs_health=[MILD, CRITICAL]},
consequent  = {loc=[HOME]}
```

In societies where vaccination rates are low, agents learn that practicing self-isolation is commendable and that making short-term concessions can help avoid major penalties.

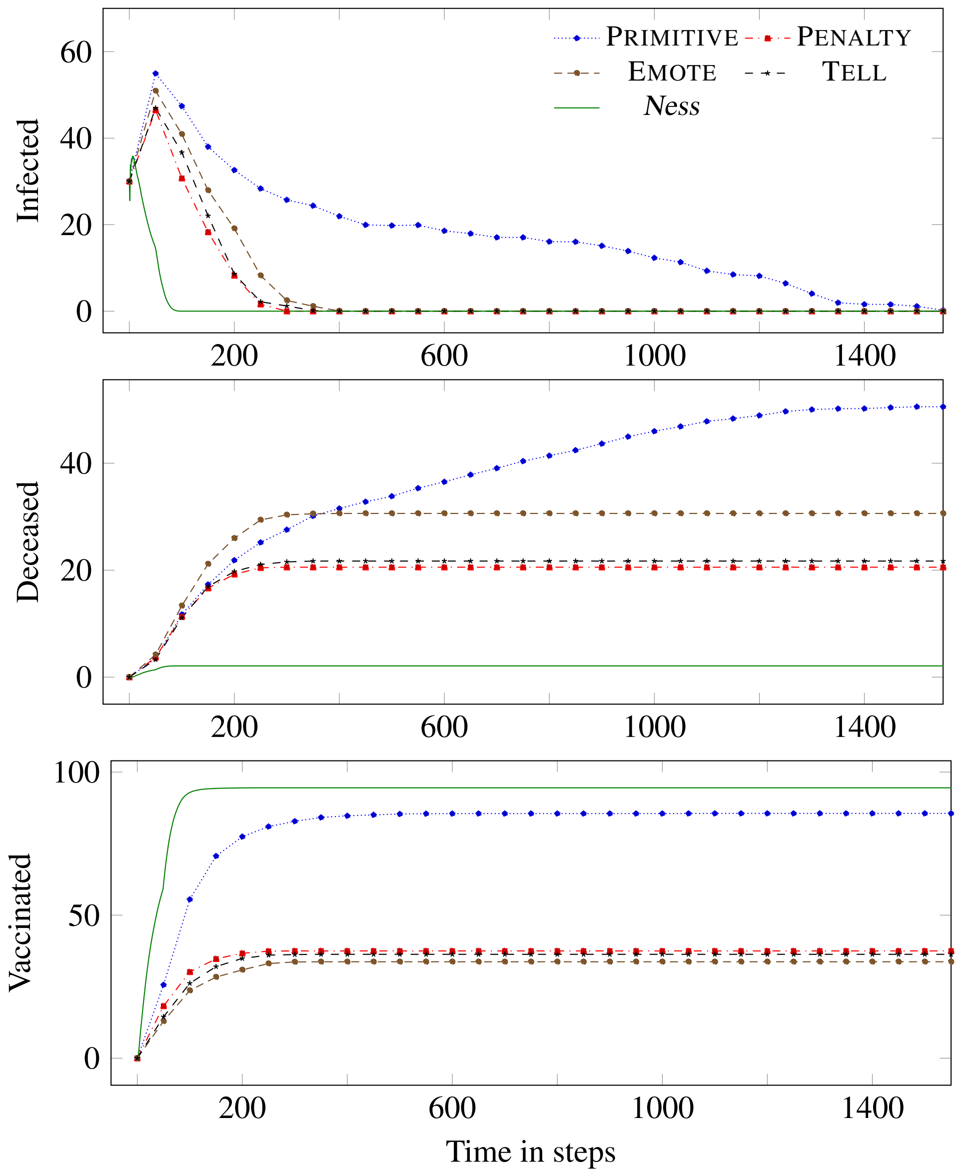


Figure 3.2: *Ness* has the least infected and deceased agents with the highest vaccination rate among all societies. However, despite a lower fraction of vaccinated agents, EMOTE has fewer infected and deceased agents, and more healthy agents than other baselines. The effect is large for the comparisons of M_{Deceased} and $M_{\text{Infections}}$. For M_{Infected}, the effect is negligible for EMOTE and small for PENALTY and TELL and large for PRIMITIVE. Appendix includes plots for the first 500 steps where the differences are noticeable.

3.4.2 $\mathrm{H}_{\mathrm{Isolation}}$

To evaluate $\mathrm{H}_{\mathrm{Isolation}}$, we measure the proportion of infected agents who stay at home ($\mathrm{M}_{\mathrm{Home}}$), the number of agents in quarantine ($\mathrm{M}_{\mathrm{Quarantine}}$), and the percentage of infected agents ($\mathrm{M}_{\mathrm{Infected}}$ in societies. Figure 3.3 exhibits plots comparing $\mathrm{M}_{\mathrm{Home}}$ and $\mathrm{M}_{\mathrm{Quarantine}}$.

We observe that *Ness* yields a higher tendency to stay isolated (0.99) when infected than the PRIMITIVE (0.61), PENALTY (0.96), EMOTE (0.95), and TELL (0.95). The effect is large for PRIMITIVE and small for the others.

Ness has a lower $\mathrm{M}_{\mathrm{Quarantine}}$ (0.00) than PENALTY (0.03), EMOTE (0.02), and TELL (0.02). The effect is small.

From $\mathrm{M}_{\mathrm{Home}}$ and $\mathrm{M}_{\mathrm{Quarantine}}$ and $\mathrm{M}_{\mathrm{Infected}}$, we observe that a norm emerges with a majority above 90% in all societies other than PRIMITIVE. Specifically, agents in societies with the mask-wearing norm learn to comply with the norm and stay self-isolated when infected. Furthermore, we see that this norm emerges fastest in *Ness* (0.99). With more subtle attitudes and information from hints, agents in *Ness* learn faster than those in TELL.

3.4.3 $\mathrm{H}_{\mathrm{Goal}}$

To evaluate $\mathrm{H}_{\mathrm{Goal}}$, we measure the goal satisfaction ($\mathrm{M}_{\mathrm{Goal}}$) in societies. Figure 3.4 plots $\mathrm{M}_{\mathrm{Goal}}$ in societies. We observe that agents in *Ness* have the highest goal satisfaction (0.31) than the PRIMITIVE (0.19), PENALTY (0.26), EMOTE (0.23), and TELL (0.26). The effect is small for EMOTE and large for the other societies.

3.5 Related Work

Research on norms and norm emergence closely relates to our contributions. Andrighetto et al. (2013) show that a combination of verbal normative information, specifically positive normative content, and negative sanction leads to higher and more stable cooperation with human subjects and agent-based simulation. These models include normative reasoning but leave out soft signals such as hints. Kalia et al. (2019) demonstrate how signals such as emotions influence norm satisfaction. Hints in *Ness* could be understood as emotions but hints serve both as an sanctioning approach and providing information about norms.

Bourgais et al. (2019) present an agent architecture that integrates cognition, contagion, personality, norms, and social relations to simulate humans and ensure explainable behaviors. Argente et al. (2022) propose an abstract normative emotional agent architecture, an extension

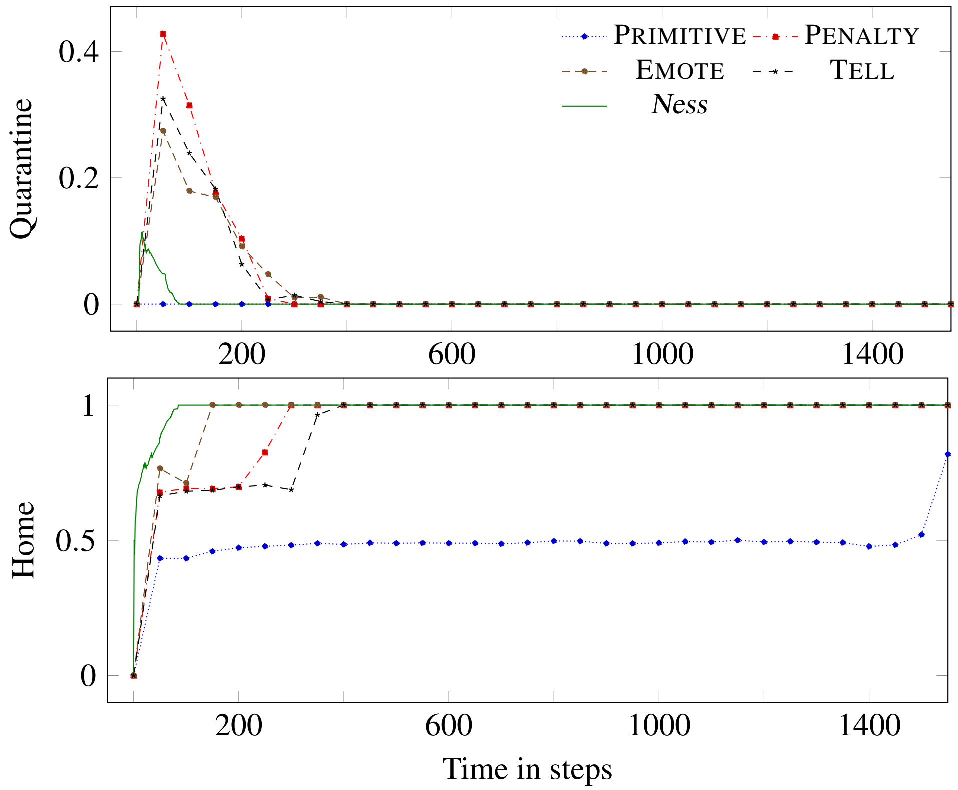


Figure 3.3: Isolation s higher in EMOTE and *Ness* i(effect is small) than in societies that lack hints. *Ness* puts fewer agents in quarantine to achieve stable cooperation than PENALTY and TELL. The effect is negligible for EMOTE and small for PENALTY and TELL with $p < 0.05$.

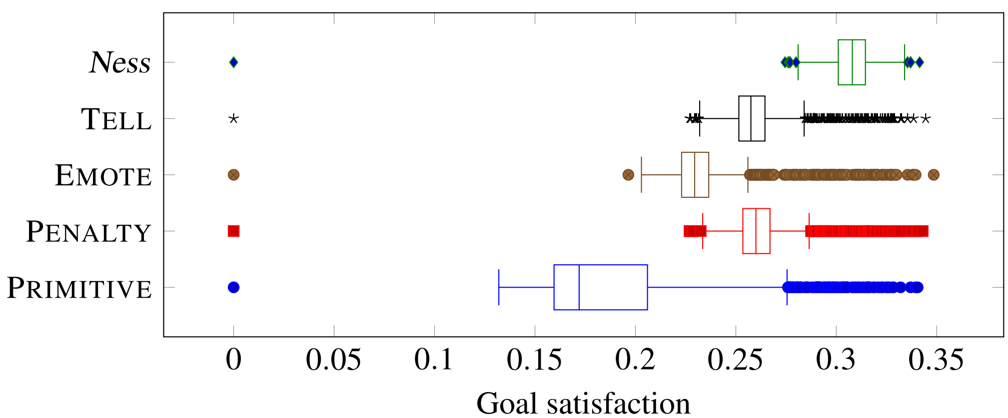


Figure 3.4: *Ness* yields more goal satisfaction than PRIMITIVE, PENALTY, EMOTE, and TELL. The effect is small for EMOTE and large for the other societies.

of BDI architecture that combines emotional, normative, and cognitive component. Tzeng et al. (2021) combine normative model, a BDI model, and emotions for the decision-making process.

Agents in *Ness* learn from their interactions with the environment and further interpret norms from various signals.

Mashayekhi et al. (2022) propose a norm emergence framework that operates on conflict detection and inequity aversion. Their framework enables agents to pass experience with utilities, associated states, and actions to agents of the same type. *Ness* agents maximize personal payoff while the social signals propel the norm emergence.

Dignum et al. (2020) associate the interventions that governments can take and their economic and social consequences with the SEIR model since effective and sustainable solutions cannot exist without considering these factors. *Ness* further takes social signaling into consideration.

de Mooij et al. (2022) develop a large-scale data-driven agent-based simulation model where each agent reasons about their internal attitudes and external factors to simulate behavioral interventions in the real world. *Ness* enables norm emergence and accommodates imposed norms.

Dell'Anna et al. (2020) introduce a norm revision component that uses data collected from interactions and an estimation of agents' preferences to modify sanctions at runtime. de Lima et al. (2019) enable agents to pick sanctions appropriate to the context. Realpe-Gómez et al. (2018) present a model in which agents incorporate personal and normative considerations. Specifically, agents make decisions that maximize their respective payoffs while appraising their group's social norms. *Ness* defines the utility function based on normative information learned from social signaling.

Airiau et al. (2014) model that supports the emergence of social norms by learning from interactions with a group of agents. In *Ness*, we include cognition via social signals. Hao et al. (2017) propose learning strategies based on local exploration and global exploration to support the emergence of social norms. Whereas their model focus on maximizing the average payoffs among agents, *Ness* focuses on investigating influence of various signals.

Morales et al. (2018) focus on the stability of synthesized norms that are verified by an evolutionary process. Savarimuthu et al. (2010) propose an algorithm to identify obligation norms based on association rule mining, a data mining technique. Pernpeintner (2021) proposes a governance approach that restrict action spaces based on publicly observable behaviors and transitions.

Levy and Griffiths (2021) propose a framework that introduces congested actions where an agent's reward is not from pairwise interaction but is a function of others' actions and the environment. *Ness* enables social learning from personal observation or by normative information sharing from explicit messages or soft signals including hints.

3.6 Discussion

During and after Covid-19, abundant research has investigated the effects of interventions against the spread of the virus. However, little research considers policy violations, which are the essential drivers of a pandemic. Modeling social signals with a framework enables a more realistic simulation of individuals' decisions, e.g., obedience or noncompliance to interventions against the spread of the pandemic.

We present an approach that combines models of social signaling to address the emergence of norms. The novelty of our approach arises from its comprehensive treatment of the three main kinds of signals that drive norm emergence: sanctions, tell, and hint. Including normative information from tell or hint enables indirect social learning, which resembles human behaviors in the real world.

3.6.1 Summary of Findings

Our main findings are that agents who signal hints and respond to normative information converge to norms faster than those who respond only to hard sanctions or explicit communication of approval or disapproval. Societies that consider hints are also robust in complying to the converged norms compared to those who do not consider hints. Specifically, in our experiments, *Ness* and EMOTE exceed the 90% of norm emergence threshold faster than other societies and their compliance to the converged norm is higher than PRIMITIVE, PENALTY, and TELL.

Our pandemic environment simulation results show that (1) *Ness* enables better control on the spread of disease than other societies, (2) agents in *Ness* and EMOTE learn the self-isolation norm faster and are more willing to isolate themselves when infected, (3) agents in *Ness* have higher goal satisfaction than the other societies. In summary, *Ness* agents effectively avoid infection risk and yield higher satisfaction than baseline agents.

3.6.2 Limitations and Threats to Validity

We made simplifying assumptions that agents can infer each other's signals and that all signals are genuine and honest. These assumptions may not apply in all cases but are essential when interacting with other AI systems or needing special care. We would assume and expect AI systems, to be honest in human-robot interaction. In addition, the signals of people who need special care reflect their needs.

3.6.3 Future Directions

While AI has been part of our daily lives nowadays, incorporating human ethics into AI becomes a necessary problem (Murukannaiah et al. 2020; Ajmeri et al. 2020; Lopez-Sanchez et al. 2017). Since human behavior is driven by the pursuit of values, studying human values helps us understand human decisions and create agents that reason over human values (Liscio et al. 2021). Social signals could also convey values. Whereas Montes and Sierra (2021) automate norm synthesis based on value promotion, an interesting direction is to embed values into autonomous agents. That is, how can we develop agents that are capable of making value-aligned decisions? A line of future research is to investigate dimensions of emotions, physical arousal, that describes the strength of the emotional state. Another future direction includes considering a mix of personality types in *Ness*. We can investigate how different values influence human interactions in future research to support high heterogeneity.

CHAPTER

4

SOCIAL VALUES ORIENTATION

4.1 Introduction

What makes people make different decisions? Schwartz (2012) defined ten fundamental human values, and each of them reflects specific motivations. Besides values, preferences define an individual's tendency to make a subjective selection among alternatives. Whereas values are relatively stable, preferences are sensitive to context and constructed when triggered (Slovic 1995).

In the real world, humans with varied weights of values evaluate the outcomes of their actions subjectively and act to maximize their utility (Schwartz 2012). In addition to values, an individual's social value orientation (SVO) influences the individual's behaviors (Van Lange 1999). Whereas values define the motivational bases of behaviors and attitudes of an individual (Schwartz 2012), social value orientation indicates an individual's preference for resource allocation between self and others (Griesinger and Livingston Jr. 1973). Specifically, social value orientation provides stable subjective weights for making decisions (Murphy and Ackermann 2014). When interacting with others is inevitable, one individual's behavior may affect another. SVO revises an individual's utility function by assigning different weights to itself and others.

Here is an example of a real-world case of SVO.

Example 4 ***SVO.***

During a pandemic, the authorities announce a mask-wearing regulation and claim that regulation would help avoid infecting others or being infected. Although Felix tests positive on the pandemic and prefers not to wear a mask, he also cares about others' health. If he stays in a room with another healthy person, Elliot, Felix will put the mask on.

An agent is an autonomous, adaptive, and goal-driven entity (Russell and Norvig 2010). Whereas many works assume agents consider the payoff of themselves, humans may further consider social preferences in the real world. e.g., payoffs of others or social welfare (Charness and Rabin 2002). When humans are in the loop along with software, there are emerging need to consider human factors when building modern software and systems. These systems should consider human values and be capable of reasoning over humans' behaviors to be realistic and trustworthy.

In a multiagent system, social norms or social expectations (Rummel 1975; Ajmeri et al. 2017) are societal principles that regulate our behavior towards one another by measuring our perceived psychological distance. Humans evaluate social norms based on human values. Most previous works related to norms do not consider human values and assume regimented environments. However, humans are capable of deliberately adhering to or violating norms. Previous works on normative agents consider human values and theories on sociality (Ajmeri et al. 2020; Verhagen 2000) in decision-making process. SVO as an agent's preference in a social context has not been fully explored.

Contributions

We investigate the following research question.

RQ_{SVO}. How do the preferences for others' rewards influence norm compliance?

To address $\mathrm{RQ}_{\mathrm{SVO}}$, we develop *Fleur*, an agent framework that considers values, personal preferences, and social norms when making decisions. Our proposed framework *Fleur* combines world model, cognitive model, emotion model, and social model. Since values are abstract and need further definition, we start with social value orientations, the stable preferences for resource allocation, in this work. Specifically, *Fleur* agents take into account social value orientation in utility calculation.

Findings

We evaluate *Fleur* via an agent simulation of a pandemic scenario designed as an iterated single-shot and intertemporal social dilemma game. We measure compliance, social experiences, and invalidation during the simulation. We find that the understanding of SVO helps agents to make more ethical decisions.

Organization

Section 4.2 presents the related works. Section 4.3 describes the schematics of *Fleur*. Section 4.4 details the simulation experiments we conduct and the results. Section 4.5 presents our conclusion and directions for future extensions.

4.2 Related Works

Griesinger and Livingston Jr. (1973) present a geometric model of SVO, the social value orientation ring as Figure 4.2. Van Lange (1999) proposes a model and interprets prosocial orientation as enhancing both joint outcomes and equality in the outcomes. Declerck and Bogaert (2008) describe social value orientation as a personality trait. Their work indicates that prosocial orientation positively correlates with adopting others' viewpoints and the ability to infer others' mental states. On the contrary, an individualistic orientation shows a negative correlation with these social skills. *Fleur* follows the concepts of social preferences from Griesinger and Livingston Jr. (1973).

Szekely et al. (2021) show that high risk promotes robust norms, which have high resistance to risk change. de Mooij et al. (2022) build a large-scale data-driven agent-based simulation model to simulate behavioral interventions among humans. Each agent reasons over their internal attitudes and external factors in this work. Ajmeri et al. (2018) show that robust norms emerge among interactions where deviating agents reveal their contexts. This work enables agents to empathize with other agents' dilemmas by revealing contexts. Instead of sharing contexts, values, or preferences, *Fleur* approximates others' payoff with observation. Serramia et al. (2018) consider shared values in a society with norms and focus on making ethical decisions that promote the values. Ajmeri et al. (2020) propose an agent framework that enables agents to aggregate the value preferences of stakeholders and make ethical decisions accordingly. This work takes other agents' values into account when making decisions. Mosca and Such (2021) describe an agent framework that aggregates the shared preferences and moral values of multiple users and makes the optimal decisions for all users. Kalia et al. (2019) investigate the relationship between norm outcomes and trust and emotions. Tzeng et al. (2021) consider emotions as sanctions. Specifically, norm satisfaction or norm violation may trigger self-directed and other-directed emotions, which further enforce social norms. Dell'Anna et al. (2019) propose a mechanism to regulate a multiagent system by revising the sanctions at runtime to achieve runtime norm enforcement.

Agrawal et al. (2022) provide and evaluate explicit norms and explanations. Winikoff et al. (2021) construct comprehensible explanations with beliefs, desires, and values. Kurtan and Yolum (2021) estimate privacy values with existing shared images in a user's social network. Tielman et al. (2019) derive norms based on values and contexts. However, these works do not consider the differences between agents and the influences of an individual's behavior on others. Mashayekhi et al. (2022) model guilt based on inequity aversion theory for an individual perspective on prosociality. In addition, they consider justice from a societal perspective on prosociality. Whereas Mashayekhi et al. (2022) assume agents may be self-interested and their decisions may be affected by others' performance, *Fleur* investigates the influence of social value orientations.

Table 4.1 summarizes related works on ethical agents. Adaptivity describes the capability of responding to different contexts. Empathy defines the ability to consider others' gain. The information share indicates information sharing among agents. The information model describes the applied models to process information and states. Among varied information models, contexts describe the situation in which an agent stands. Emotions are the responses to internal or external events or objects. Guilt is an aversive self-directed emotion. Explicit norms state causal normative information, including antecedents and consequences. Values and preferences both define desirable or undesirable states.

4.3 *Fleur*

We now discuss the schematics of *Fleur* agents.

Figure 4.1 shows the architecture of *Fleur*. *Fleur* agents consists of five main components: cognitive model, emotion model, world model, social model, and a decision module.

4.3.1 Cognitive Model

Cognition relates to conscious intellectual activities, such as thinking, reasoning, or remembering, among which human values and preferences are essential. Specifically, values and preferences may change how an individual evaluates an agent, an event, or an object. In *Fleur*,

Table 4.1: Comparisons of works on ethical agents with norms and values.

Research	Adaptivity	Empathy	Information Share	Information Model
Fleur	✓	✓	✗	Preferences & Emotions & Contexts
Agrawal et al. (2022)	✓	✗	✓	Explicit norms
Ajmeri et al. (2018)	✓	✓	✓	Contexts
Ajmeri et al. (2020)	✓	✓	✓	Values & Value preference & Contexts
Kalia et al. (2019)	✓	✗	✗	Trust & Emotions
Kurtan and Yolum (2021)	✓	✗	✗	Values
Mashayekhi et al. (2022)	✓	✓	✓	Guilt
Mosca and Such (2021)	✓	✓	✓	Preferences & Values
Serramia et al. (2018)	✓	✗	✗	Values
Tielman et al. (2019)	✓	✗	✓	Values & Contexts
Tzeng et al. (2021)	✗	✗	✗	Emotions
Winikoff et al. (2021)	✓	✗	✗	Values & Beliefs & Goals

we start with the preferences of individuals in the allocation of resources. Preferences are the attitudes toward a set of objects in psychology (Slovic 1995). For instance, SVO provides agents with different preferences over resource allocations between themselves and others. Figure 4.2 demonstrates the reward distribution of different SVO types. The horizontal axis measures the resources allocated to oneself, and the vertical axis measures the resources allocated to others. Let $\overrightarrow{R} = (r_1, r_2, \ldots, r_n)$ represent the reward vector for a group of agents with size n. The reward for agent i considering social aspect is:

$$reward_i = r_i \cdot \cos\theta + r_{-i} \cdot \sin\theta \tag{4.1}$$

where r_i represents the reward for agent i and r_{-i} is the mean reward of all other agents interacting with agent i. Here we adopt the reward angle from McKee et al. (2020) and represent agents' social value orientation with θ. We define $\theta \in \{90°, 45°, 0°, -45°\}$ as SVO $\in$ {altruistic, prosocial, individualistic, competitive}, respectively. With the weights provided by SVO, the presented equation enables the accommodation of social preferences.

In utility calculation, we consider two components: (1) extrinsic reward and (2) intrinsic

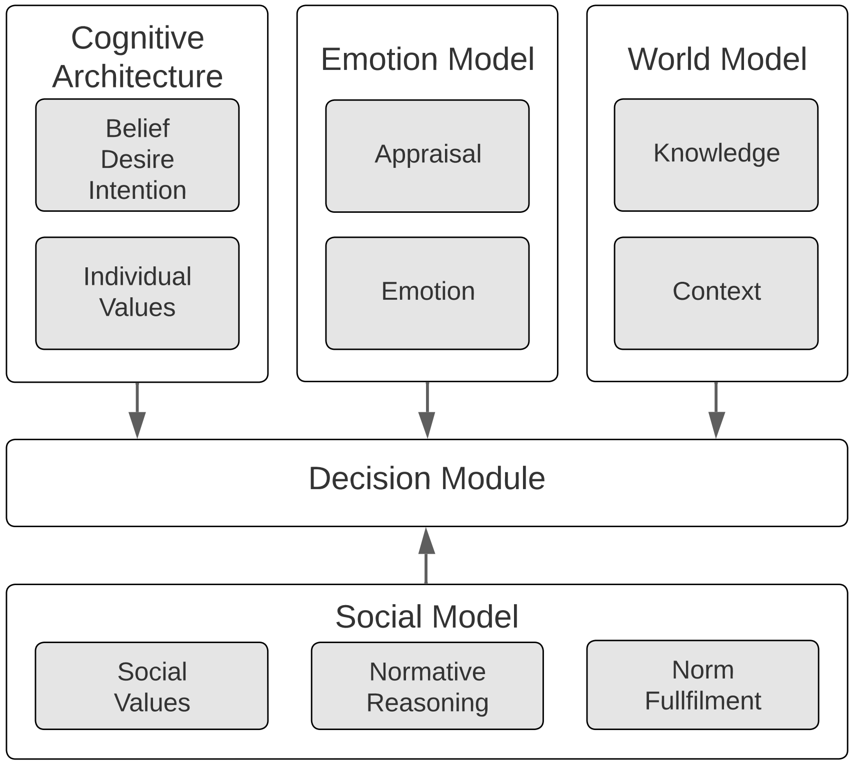


Figure 4.1: *Fleur* architecture.

reward. Whereas extrinsic rewards come from the environment, intrinsic rewards stem from internal stats, e.g., human values and preferences.

We extend the Belief-Desire-Intention (BDI) architecture (Rao and Georgeff 1991). An agent forms beliefs based on the information from the environment. The desire of an agent represents having dispositions to act. An agent's intention is a plan or action to achieve a selected desire.

Take Example 4 for instance. Since Felix has an intention to maximize the joint gain with Elliot, he may choose a strategy to not increase his payoff at the cost of others' sacrifice.

4.3.2 Emotion Model

We adopt the OCC model of emotions (Ortony et al. 1988). Specifically, our emotion model appraises an object, an action, or an event and then triggers emotions. We consider emotional valence and assume norm satisfaction or norm violation yields positive or negative emotions if

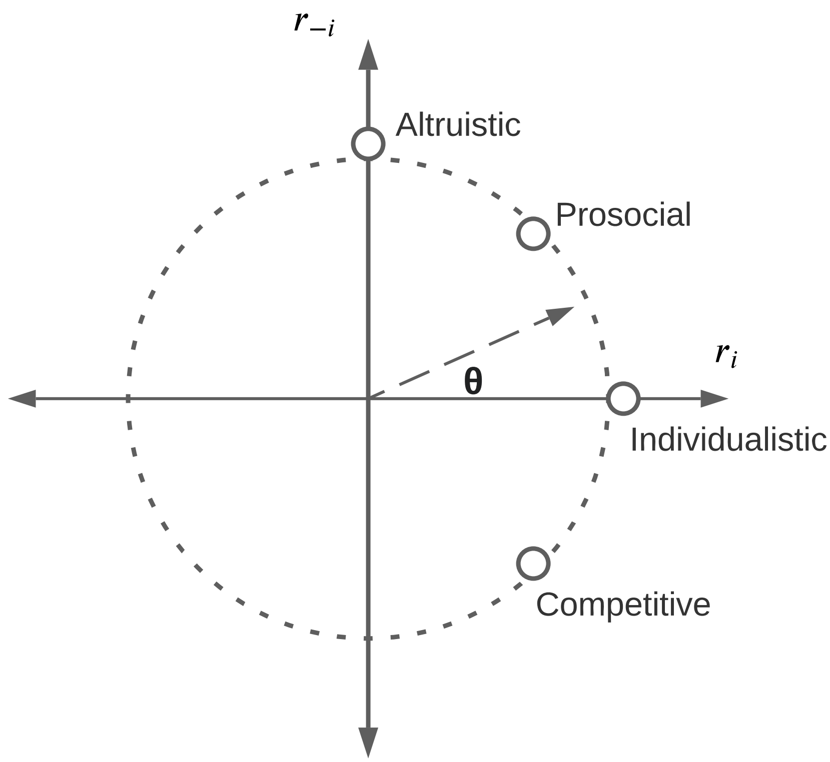


Figure 4.2: Representation of Social Value Orientation (Griesinger and Livingston Jr. 1973; McKee et al. 2020). r_i denotes outcome for oneself and r_{-i} denotes outcomes for others.

self behaviors align with the norms.

4.3.3 World Model

The world model describes the contexts in which *Fleur* agents stand and represents the general knowledge *Fleur* agents possess. A context is a scenario that an agent faces. Knowledge in this model are facts of the world. In Example 4, the context is that an infected individual, Felix, seeks to maximize the collective gain of himself and a healthy individual, Elliot. In the meantime, Felix acknowledges that a pandemic is ongoing.

4.3.4 Social Model

The social model of an agent includes social values, normative reasoning, and norm fulfillment. Social values refer to the values of a society, while individual values delineate the values that characterize an individual. Agents use the normative-reasoning component to reason over states, norms, and possible outcomes of satisfying or violating norms. Norm fulfillment checks if a norm has been fulfilled or violated with the selected action. Sanctions may come after norm fulfillments or violations.

4.3.5 Decision Module

The decision module selects actions based on agents' payoffs and individual values. We apply Q-Learning (Watkins and Dayan 1992), a model-free reinforcement learning algorithm that learns from trial and error, to our agents. Q-Learning approximates the action-state value $Q(s,a)$ (Q value), with each state and action:

$$Q'(s,a) = Q(s,a) + \alpha \times (r_t + \gamma \max_{a'} Q(s',a') - Q(s,a)) \tag{4.2}$$

where $Q'(s,a)$ represents the updated Q-value after performing action a at state s. s' represents the next state and a' refers to possible actions in s'. α denotes the learning rate in the Q-value update function, and r_t represents the rewards received at state s after acting a. γ defines the reward discount rate, which characterizes the importance of future rewards. Agents observe the environment, form their beliefs about the world, and update their state-value with rewards via interactions. By approximating the action-state value, the Q-Learning algorithm finds the optimal policy via the expected and cumulative rewards.

Algorithm 3 describes the agent interaction in our simulation.

4.4 Experiments

We now describe our experiments and discuss the results.

4.4.1 Experimental Scenario: Pandemic Mask Regulation

We build a pandemic scenario as an iterated single-shot and intertemporal social dilemma. We assume that the authorities have announced a masking regulation. In each game, each agent selects from the following two actions: (1) wear a mask, and (2) not wear a mask. Each agent

Algorithm 3: Decision loop of a *Fleur* agent

```
Initialize one agent with its desires D and preference P and SVO angle θ;
Initialize action-value function Q with random weights w;
for t=1,T do
    Pair up with another agent pn to interact with;
    Observe the environment (including the partner and its θ) and form beliefs b_t;
    With a probability ε select a random action a_t
      Otherwise select a_t = argmax_a Q(b_t, a; w)
    Execute action a_t and observe reward r_t;
    Observe the environment (including the partner) and form beliefs b_{t+1};
    Activate norms N with beliefs b_t, b_{t+1}, and action a_t;
    if N != ∅ then
        Sanction the partner based on a_t and its behavior;
    end
end
```

has its inherent preferences and social value orientation. An agent forms a belief about its partner's health based on its observation. During the interaction, the decision an agent makes affects itself and others. The collective behaviors among agents determine the the dynamics in a society. Each agent receives the final points from its own action and effects from others: $R_{sum} = P_{i\_self} + P_{i\_other} + S_j$. P_{i_self} denotes the payoff from the action that agent i selects considering the reward distribution in Figure 4.2 and self-directed emotions. P_{i_other} is the payoff from the action that the other agent performs. S_j denotes the other-directed emotions from others towards agent i.

Table 4.2: Payoff for an actor and its partner based on how the actor acts and how its action influence others. Column Actors show the points from the actions of the actor. Column Partners display the points from the actions to the partner.

Health		Actions			
		Mask		No mask	
Actor	Partner	Actor	Partner	Actor	Partner
healthy	healthy	0.00	0.00	0.00	0.00
healthy	infected	1.00	0.00	−1.00	0.00
infected	healthy	0.00	1.00	0.00	−1.00
infected	infected	0.50	0.50	−0.50	−0.50

Table 4.3: Payoff for decisions on preferences.

Type	Decisions	
	Satisfy	Dissatisfy
Preference	0.50	0.00

Table 4.4: Payoff for decisions on norms.

Actor	Partner	
	Wear	Not-Wear
Wear	0.10	-0.10
Not-Wear	0.00	0.10

4.4.2 Experimental Setup

We develop a simulation using Mesa (Masad and Kazil 2015), an agent-based modeling framework in Python for creating, visualizing, and analyzing agent-based models. We ran the simulations on a device with 32 GB RAM and GPU NVIDIA GTX 1070 Ti.

We evaluated *Fleur* via a simulated pandemic scenario where agents' behaviors influence the collective outcome of the social game. A game-theoretical setting may be ideal for validating the social dilemma with SVO and norms. However, real-world cases are usually non-zero-sum games where one's gain does not always lead to others' loss. In our scenario, depending on the context, the same action may lead to different consequences for the agent itself and its partner. For instance, when an agent is healthy and its partner is infected, wearing a mask gives the agent a positive payoff from the protection of the mask but no payoff for its partner. Conversely, not wearing a mask leads to a negative payoff for the agent and no payoff for its partner. The payoff given to the agent and its partner corresponds to the X and Y axis in Figure 4.2. When formalizing social interactions with SVO in game-theoretical settings, the payoffs of actions for an agent and others are required information.

We incorporated beliefs and desires, and intentions into our agents. An agent observes its environment and processes its perception, and forms its beliefs about the world. In each episode, agents pair up to interact with one another and sanction based on their and partners' decisions (Table 4.4).

Context. A context is composed of attributes from an agent and others and the environment as shown in Table 4.2. We frame the simulation as a non-zero-sum game where one's gain does not necessarily lead to the other parties' loss.

Preference. In psychology, preferences refer to an agent's attitudes towards a set of objects. In our simulation, we set 40% of agents to prefer to wear and prefer not to masks individually. The rest of the agents have a neutral attitude on masks. The payoffs for following the preferences are listed in Table 4.3.

Social Value Orientation. Social value orientation defines an agent's preference for allocating resources between itself and others. We consider altruistic, prosocial, individualistic, and competitive orientations selected from Figure 4.2.

4.4.3 Hypotheses and Metrics

We compute the following measures to address our research question RQ_{SVO}.

Compliance The percentage of agents who satisfy norms

Social Experience The total payoff of the agents in a society

Invalidation The percentage of agents who do not meet their preferences in a society

To answer our research question RQ_{SVO}, we evaluate three hypotheses that correspond to the specific metric, respectively.

$H_{Compliance}$: Preferences for others' rewards positively affect norm compliance with prosocial norms

$H_{Social\ Experience}$: The distribution of preferences for others' rewards positively affect social experiences in a society

$H_{Invalidation}$: Preferences for others' rewards negatively affect the tendency to meet personal preferences

4.4.4 Experiments

We ran a population of $N = 40$ agents in which we equally distributed our targeted SVO types: altruistic, prosocial, individualistic, and competitive. Since each game is a single-shot social dilemma, we consider each game as an episode. The training last for 500,000 steps. In evaluation, we run 100 episodes and compute the mean values to minimize deviation from coincidence. We define our five societies as below.

Mixed society A society of agents with mixed social value orientation distribution

Altruistic society A society of agents who make decisions based on altruistic concerns

Prosocial society A society of agents who make decisions based on prosocial concerns

Selfish society A society of agents who make decisions based on selfish concerns

Competitive society A society of agents who make decisions based on competitive concerns

We assume all agents are aware of a mask-wearing norm. Agents who satisfy the norm receive positive emotions from themselves and others, as in Table 4.4. Conversely, norm violators receive negative emotions. Table 4.5 summarizes results of our simulation.

Table 4.5: Comparing agent societies with different social value orientation distribution on various metrics and their statistical analysis with Glass' Δ and p-value. Each metric row shows the numeric value of the metric after simulation convergence.

		Compliance	Social Experience	Invalidation
S_{mixed}	$\bar{X}$	63.40%	0.45	29.60%
	p-value	–	–	–
	Δ	–	–	–
$S_{altruistic}$	$\bar{X}$	69.70%	0.55	**33.40%**
	p-value	< 0.001	< 0.001	< 0.001
	Δ	0.66	0.61	0.46
$S_{prosocial}$	$\bar{X}$	**70.25%**	**0.57**	32.28%
	p-value	< 0.001	< 0.001	< 0.05
	Δ	0.72	0.68	0.33
$S_{selfish}$	$\bar{X}$	65.10%	0.47	26.90%
	p-value	0.22	0.42	< 0.05
	Δ	0.18	0.12	0.33
$S_{competitive}$	$\bar{X}$	54.08%	0.22	28.88%
	p-value	< 0.001	< 0.001	0.54
	Δ	0.98	1.31	0.09

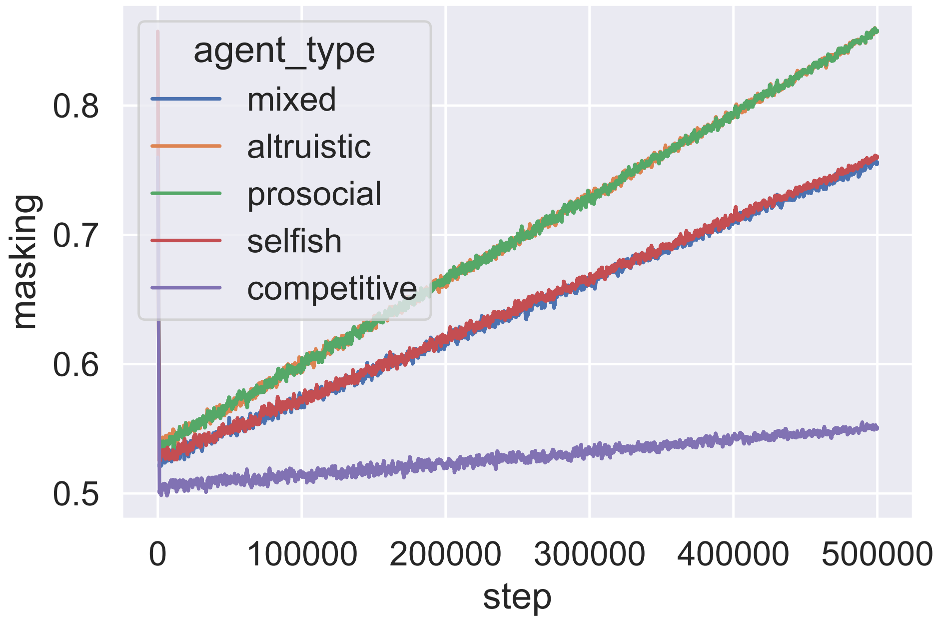


Figure 4.3: Compliance in training phase: The percentage of norm satisfaction in a society.

Figure 4.3 displays the compliance, the percentage of agents who satisfy norms, in the mixed and baseline-agent societies. We find that the compliance in the altruistic and prosocial-agent society, averaging at 69.70% and 70.25%, is higher than in the mixed (63.40%) and agent societies have no positive weights on others' payoff (65.10% and 54.08% for selfish and competitive-agent societies, respectively). The differences in the results of altruistic and prosocial-agent societies are statistically significant with medium effect ($p < 0.001$; Glass' $\Delta > 0.5$). Conversely, the competitive-agent society has the least compliance, averaging at 54.08%, with $p < 0.001$ and Glass' $\Delta > 0.8$. The results of the selfish-agent society (65.10%) shows no significant difference with $p > 0.05$ and Glass' $\Delta \approx 0.2$.

There are 25% of agents in the mixed-agent society are competitive agents. Specifically, they prefer to minimize others' payoff. A competitive infected agent may choose not to wear a mask when interacting with other healthy agents in this scenario. In the meantime, the selfish agents would maximize their self utility without considering others. Therefore, the behaviors of selfish and competitive agents may decrease compliance in the mixed-agent society.

Figure 4.4 compares the average payoff in the mixed and baseline-agent societies. The social experience in the altruistic and prosocial-agent society, averaging at 0.55 and 0.57, is higher than in the mixed (0.45) and agent societies have no positive weights on others' payoff (0.47 and 0.22 for selfish and competitive-agent societies, respectively). The differences in

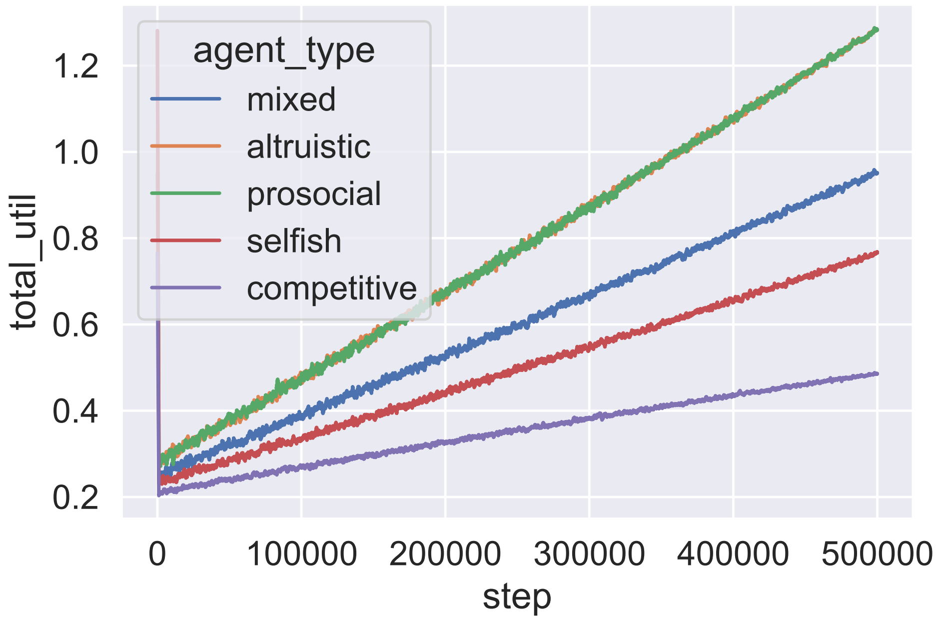


Figure 4.4: Social Experience in training phase: The total payoff of the agents in a society.

the results of altruistic and prosocial-agent societies are statistically significant with medium effect ($p < 0.001$; Glass' $\Delta > 0.5$). On the contrary, the competitive-agent society has the least social experience, averaging at 0.22, with $p < 0.001$ and Glass' $\Delta > 0.8$. The results of the selfish-agent society (0.47) shows no significant difference with $p > 0.05$ and Glass' $\Delta < 0.2$.

The mixed-agent society shows similar results as the selfish-agent society. Although 50% of the mixed-agent society agents are altruistic and prosocial, the competitive agents would choose to minimize others' payoff without hurting their self-interests.

Since the selfish agents do not care about others, they would act for the sake of their benefit. The selfish and competitive behaviors diminish the social experiences in society.

Figure 4.5 compares invalidation, the percentage of agents who do not meet their preferences in the mixed and baseline-agent societies.

The invalidation in the altruistic and prosocial-agent society, averaging at 33.40% and 32.28%, is higher than in the mixed (29.60%) and agent societies have no positive weights on others' payoff (26.90% and 28.88% for selfish and competitive-agent societies, respectively). The differences in the results of altruistic and prosocial-agent societies are statistically significant with small or medium effect ($p < 0.001$; Glass' $\Delta > 0.2$). On the contrary, the selfish-agent society has the least invalidation, average at 26.90%, with $p < 0.05$ and Glass' $\Delta > 0.2$. The results of the competitive-agent society (28.88%) shows no significant difference with $p > 0.05$

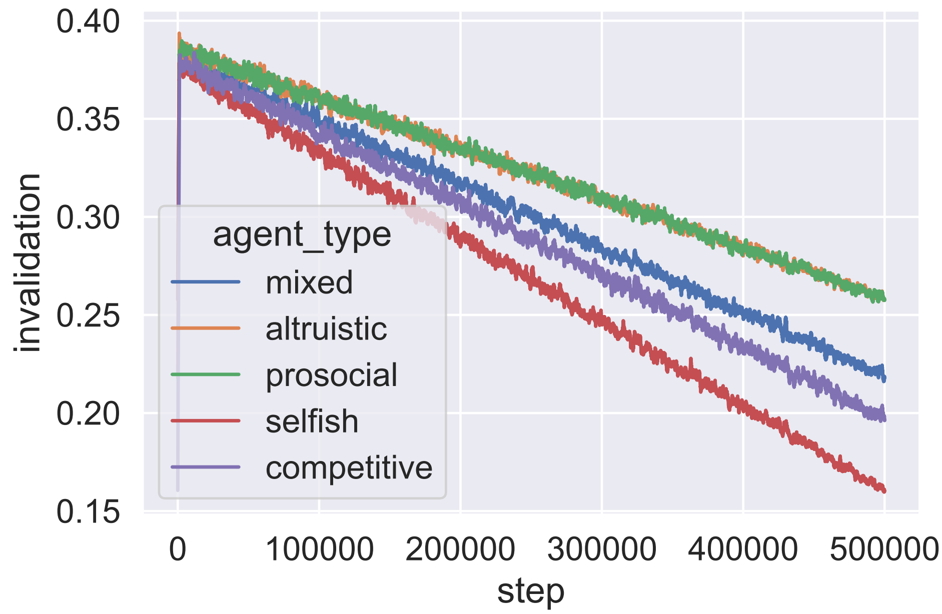


Figure 4.5: Invalidation in training phase: The percentage of agents who do not meet their preferences in a society.

and Glass' $\Delta < 0.2$.

While agents who consider others' rewards positively achieve better compliance and social experiences, these achievements are based on their sacrifice of preferences. The altruistic and prosocial agent societies have the most percentage of agents who do not meet their preferences.

4.4.5 Threats to Validity

First, our simulation has a limited action space. Moreover, different actions may have the same payoff in some contexts. Other behaviors may better describe different types of SVO, yet our focus is on showing how SVO influences normative decisions.

Second, we represent actual societies as simulations. While differences in preference and SVO among people are inevitable, we focus on validating the influence of SVO.

Third, to simplify the simulation, we assume fixed interaction, whereas real-world interactions tend to be random. An agent may interact with one another in the same place many times or have no interaction. We randomly pair up all agents to mitigate this threat and average out the results.

4.5 Conclusions and Directions

We present an agent architecture that integrates cognitive architecture, world model, and social model to investigate how social value orientation influences compliance with norms. We simulate a pandemic scenario in which agents make decisions based on their individual and social preferences. The simulations show that altruistic and prosocial-agent societies comply better with the mask norm and bring out higher social experiences. However, altruistic and prosocial agents trade their personal preferences for compliance and social experiences. The results between the mixed and selfish-agent societies show no considerable difference. The competitive agents in the mixed-agent society may take the responsibility.

Future Directions

Our possible extensions include investigating an unequal distribution of SVO in *Fleur* and applying real-world data in the simulation. Other future directions are incorporating values into agents, and revealing adequate information to explain and convince others of inevitable normative deviations (Agrawal et al. 2022; Murukannaiah et al. 2020; Woodgate and Ajmeri 2022).

CHAPTER

5

DECISION AND RATIONALE WITH VALUES

5.1 Introduction

Social norms define the shared standard of acceptable behaviors in a society (Von Wright 1963). An agent forms legitimate expectations of the behavior of others based on norms. An agent may deviate from a norm in exercising its autonomy (Singh and Singh 2023). A deviation from a norm may lead to a positive or a negative sanction. A deviation from a social norm may be excused with an acceptable rationale. Further, to be trusted by humans, an agent should be able to provide rationales for their decisions (Winikoff et al. 2021; Ayci et al. 2023).

Example 5 ***Sharing rationale.*** *Alice comes to office with a mask where she notices Bella not wearing a mask. Bella justifies her decision by stating that, first, there is no mask mandate in the office as the surrounding environment is safe. Second, she hates wearing a mask because wearing one causes her eczema. Alice agrees with Bella's perspective.*

Langley (2019) and Miller (2019) suggest that the goal of rationales is to provide necessary information for a decision. In practice, rationale include additional information that others may be unable to observe, e.g., beliefs and preferences. Crafting a rationale is challenging. Rationales may be verbose, leading to information overload. They may include private information, which one may not prefer sharing.

Example 6 ***Adapting rationale.*** *Bella and Alice both value health. Although Bella prefers not to wear a mask due to skin issues, sharing that may not be necessary. Bella justifies her behavior of not wearing a mask by stating that the surrounding environment is safe and that a mask is not needed. Alice finds Bella's rationale convincing.*

A good decision made by an agent should go beyond physical gain and be aligned with human values (Woodgate and Ajmeri 2022; Yazdanpanah et al. 2023). Whereas values guide motivations and drive decisions, rationales or information aligned with values best justify one's behaviors (Liscio et al. 2023; Winikoff et al. 2021).

We posit that a good rationale is the one which (1) considers the context in which the decision is made by an agent and (2) considers the values, understood as motivational bases of one's behaviour (Schwartz 2012), of the decision-maker (the agent who produces a rationale) and the receiver of the rationale. Instead of sharing all available information in a rationale, it is beneficial for an agent to share only relevant pieces of information which align well with self and others' values. Sharing only relevant pieces preserves privacy of the provider of the rationale and ensures that the receiver of a rationale is not overwhelmed by unnecessary information.

Values reflect various concerns in decision-making and conflict resolution. Whereas value-aligned rationales enable agents to justify their behavior, deliberating over others' values may increase convincingness and acceptance. Based on the preceding intuition, we make the following contribution.

Contribution We create *Exanna*, a framework that incorporates values in decision-making, rationale generation, and reasoning over rationale. Whereas other works focus on making agent decisions interpretable to humans, *Exanna* agents provide rationales to both agents and humans.

Findings We evaluate *Exanna* via a multiagent simulation study considering a pandemic scenario. We consider agent societies with three characteristics for producing rationales: share all, share decision rules, and share value-aligned rules. With *Exanna*, we find that agents who consider value preferences when giving rationales achieve higher conflict resolution. Also, rationales aligned with values but with less private information lead to better social experience.

Novelty This work presents essential perspectives on decisions and rationales with values. First, an individual's decision-making and evaluation involves values, but values with higher weights dominate the decision. A compelling rationale often states causal relationships with the esteemed values of agents. An *Exanna* agent makes decisions based on its or its stakeholder's values. Upon generating rationales, the agent discloses the causal effects based on values, which are interpretable to both agents and humans. Second, individuals evaluate actions or states based on their values. *Exanna* incorporates values in state evaluation.

Organization Section 5.2 discusses relevant related works. Section 5.3 details the *Exanna* framework. Section 5.4 describes a simulated pandemic scenario for evaluation and the results. Section 5.6 concludes with listing potential future directions.

5.2 Related Work

Research on agents with rationales and values is related.

5.2.1 Agents and Rationales

Hind et al. (2019) leverage existing supervised machine-learning techniques to generate rationales together with decisions without values involved and without exposing the inner details of the model. Whereas Hind et al. generate rationales based on the rationales in the training set, *Exanna* generates rationales based on context and values.

Georgara et al. (2022) propose an algorithm that wraps up any team formation algorithm to build justifications on why specific teams are formed. Specifically, Georgara et al. build justifications based on contrastive explanations and by exploring what-if scenarios. A causal attribution explains why a behavior occur. We provide causal attribution of the selected action, precisely the premise, as rationales and wrap the rationales with values.

Wang et al. (2021) formulate rationales with the simplest subset of features with the proposed search algorithm. This algorithm finds sufficient rationales by modifying the beam search algorithm and leveraging the tractability of expected predictions. The found set of features is sufficient as causal attribution for probabilistic solid guarantees on model behavior under observed data distribution. Contreras et al. (2022) propose a mirror model and assume a high understandability from performing similar to an observer's mental simulation. They apply deep Q-network and saliency maps in rationale generation, highlighting related input features as rationales. These works reveal each feature related to the model behavior.

Ajmeri et al. (2018) propose Poros, a framework that considers no values and shares full context as a rationale. Therefore, agents can make decisions from the perspective of others. In *Exanna*, agents *selectively* share information based on its and others' *values*.

5.2.2 Agents and Values

Lera-Leri et al. (2022) propose a method that considers a range of ethical principles from maximum utility to maximum fairness, for the aggregation of value systems instead of one single value. Ajmeri et al. (2020) present a framework that aggregates the value preferences of users to make ethically appropriate decisions. In addition to making decisions based on values, *Exanna* agents justify their behaviors and evaluate rationales based on their *values*.

Mosca and Such (2021) propose an agent that supports values in multiuser settings via generating optimal policy considering the preferences and values of users. They justify solutions through contrastive explanations and positive answers. The causal attribution includes (1) a suggested action and the inputs of all users and (2) possible consequences from the user's preference. Liao et al. (2023) propose an ethical recommendation component involving multiple stakeholders, which employs methods from normative systems and formal argumentation to achieve agreements among agents. Whereas these works generate explanations to present causal attribution with all the necessary information, *Exanna* further wraps rationales with values.

Agrawal et al. (2022) propose an agent that shares norms as causal attributions, and considers no values. Specifically, each agent evolves and learns rules of optimal behaviors with no values involved. *Exanna adaptively* shares learned rules as rationales that align with individual's *values*.

Montes and Sierra (2022) propose a methodology to synthesize parametric normative systems based on value promotion. Whereas Montes and Sierra focus on the design of moral norms synthesis and do not consider internal reasoning on norms, *Exanna* focuses on internal reasoning and justifying behavior based on values. Ogunniye and Kökciyan (2023) propose an ontology to represent the privacy domain that includes norms for social contexts, privacy preferences, and privacy policies. Ogunniye and Kökciyan introduce an argumentation-based dialogue to provide justifications during multi-party dialogues. In addition, the dialogue helps agents to reason about contextual norms and resolve privacy conflicts among agents. Di Scala and Yolum (2023) propose a new privacy agent for content concealment, equity of treatment, the collaboration of users, and the rationalization of actions. Specifically, Scala and Yolum provide textual output by considering outcomes and providing feedback to the user, such as a summary or detailed advice on possible actions to improve performance. Whereas these works claim that argumentation-based dialogue facilitates the exchange of rationales, *Exanna* provides a value-centered rationale for the made decision.

Table 5.1 summarizes the above comparisons emphasizing values and rationales.

Table 5.1: Summary of comparisons with related work with respect to the application of values in decisions and rationales (generation and evaluation). The ✓ and ✗ notions signify that values are applied and not applied, respectively.

	Rationale	Values applied in		Rationale representation
		Decision	**Rationale**	
Lera-Leri et al. (2022)	✗	✓	✗	No rationales provided
Ajmeri et al. (2020)	✗	✓	✗	No rationales provided
Agrawal et al. (2022)	✓	✗	✗	Norm as causal attribution but no information hiding
Contreras et al. (2022)	✓	✗	✗	Highlighted input features in deep Q-network but no information hiding
Wang et al. (2021)	✓	✗	✗	The prediction and a minimum subset of inputs but no information hiding
Hind et al. (2019)	✓	✗	✗	Texts predicted via supervised learning, along with the predicted action
Ajmeri et al. (2018)	✓	✗	✗	Full context
Di Scala and Yolum (2023)	✓	✓	✓	Outcomes or advice based on complete information but no information hiding
Ogunniye and Kökciyan (2023)	✓	✓	✓	A sequence of communications but no information hiding
Mosca and Such (2021)	✓	✓	✓	Suggested action based on the inputs from all users, along with the possible outcome of the user's preference as causal attribution but no information hiding
Exanna	✓	✓	✓	Behavior rules (with information hiding) and alignment with values

5.3 Method

We now describe the schematics and decision making in *Exanna* along with its rationale components.

5.3.1 Schematics of an *Exanna* Agent

Belief is an agent's interpretation of the world, which is formed based on its observations. b_t indicates the belief from observation at time t. *Exanna* agents store beliefs as pairs of attributes and their bindings.

Context is the information that characterizes the situation of an agent. Context is represented as a set of attribute-binding pairs. An example of context is as below.

```
{Risk=None, Preference= ¬Wear, InteractWith=Colleague,
OtherAgentType=Health, RiskFromAnother=High,
OtherAgentPreference=Wear, Location=Office}
```

A context comprises public (e.g., an agent's location) and private (e.g., beliefs, preferences, and values) attributes.

Goal defines the desired state that an agent wants to achieve. The outcome of a goal has a binary value, indicating whether the goal is achieved or not after performing the selected action

Action is the methodology to change the state and, therefore, approach the goals. We represent an action as a where $a \in A$ and A is the set of available actions.

Preference refers to a subjective inclination for an option over other alternatives. We represent a preference as p and $p \in O$ where O is the option space.

Decision rule is the mapping between an observation and a reasonable action, represented as if-then logic. A decision rule includes a premise and a consequent. The premise of a decision rule is a set of attribute-binding pairs. The consequent of a decision rule is an action to be taken when the premise holds. An example rule is

```
{Risk=None, InteractWith=Colleague} => ¬Wear
```

Norm is the expected behavior or the behavior of the majority in a group. When a majority applies the same decision rule, the rule becomes a norm. In *Exanna*, a norm uses the same if-then representation as a decision rule.

Sanction is the response to norm violation or satisfaction. A sanction can be a positive, negative, or neutral reaction from one agent to another.

Payoff refers to the outcome or result an agent receives in a given state after taking an action.

Values refers to motivational goals of agents. A subset of values is applicable within a context (Liscio et al. 2021) and agents have a preference order over those values.

Value preferences is a preference order over various values for one context. We store each value preference $V_{context}$ in a tuple where numbers add up to 1. v_i denotes the weight of one value in one value preference ($v_i \in V_{context}$) where $0 \leq v_i \leq 1$ and $\sum_{i=1}^{n} v_i = 1$. We treat each $\langle V_{context} \rangle$ as an attribute and store the corresponding preferences as its binding. For instance, an agent with value preferences $V = \{V_{pandemic} = \{v_{health} = 0.6,\ v_{privacy} = 0.4\}; V_{normal} = \{v_{health} = 0.4,\ v_{privacy} = 0.6\}\}$ indicates that the agent values health over privacy during a pandemic but the opposite in a normal context.

5.3.2 Interaction and Decision Making

Interactions in *Exanna* are between an actor, the rationale giver, and an observer agent, the rationale receiver. An actor agent selects an action based on its goal. Based on the chosen action, the actor agent provides a rationale to the observer who witnesses its behavior. Upon receiving a rationale from the actor agent, the observer agent evaluates the rationale by making an analogous decision. With a weighted sum of payoffs, we incorporate values in decision-making where a substantial value casts a more significant effect on the final decision.

Algorithm 4 describes the pseudo-code of an agent's decision-making loop. An agent forms beliefs b_t about the world based on its observations (Line 4). An agent's payoff is a weighted sum of payoffs corresponding to an agent's values. The Q function in Line 6 and reward in Line 7 refers to the payoff calculation in Appendix .2.3. The Q function, in addition, includes feedback from others. In Line 6, the agent selects the action that gives the best payoff for b_t. If the agent interacts with another agent, for its action the agent creates rationales based on b_t and the selected action (Line 9 with Algorithm 5) and sends those rationales. Other agents who observe the action and receive the rationales evaluate the rationales (Algorithm 6) with their context and give sanctions.

5.3.3 Rationale Generation

Rationale generation in *Exanna* follows a *rule learning* process—a process of evolving rules from datasets or interactions. The basic form of a rule is an if-then expression, e.g., if premise then consequent, where the consequent holds whenever the premise is true. We adapt XCS (Butz and Wilson 2000), a rule-based learning algorithm that utilizes a genetic algorithm and reinforcement learning, which evolves a set of rules or strategies based on payoffs or rewards

Algorithm 4: Decision-making for an *Exanna* agent

Initialize agent (including value preferences V and other mental states);
Initialize rule-value function Q;
for t=*1*,T **do**
  Form beliefs b_t based on perceived state;
  Identify available actions A;
  With a probability ε select a random action $a_{actor} \in A$
   Otherwise select $a_{actor} = argmax_a Q(b_t, a)$;
  Execute action a_{actor} and observe reward r_t;
  if *Any observer agent pa* **then**
    `/* Generate rationales based on selected action and beliefs */`
    Rat = GenRationale(b_t, a_{actor});
    Send Rat to pa;
    Observe agent pa's action a_{actor};
    if *Receive rationales* Rat_{actor} *from agent pa* **then**
      Update beliefs b_t based on Rat_{actor};
    end
    `/* Generate sanctions based on beliefs and given rationales */`
    $sanction_{actor}$ = EvalRationale(Rat_{actor} if any, a_{actor}, b_t);
  end
  `/* Agents learn from reward and sanction */`
  learn(b_t, a_{actor}, $r_t + sanction_{actor}$, b_{t+1});
end

produced by the proposed actions. Unlike other machine learning techniques, XCS, in addition to generating a decision, generates a set of rules describing its decision. XCS process enables flexibility for implementation of norms and supports interpretability for humans with logical rules.

An example rule of Example 6 is

```
{Risk=None, InteractWith=Colleague} => ¬Wear
```

The premise of a learned rule is a conjunction of attribute-binding pairs, e.g., {Risk=None, InteractWith=colleague}. The consequent of a learned rule is an action to be taken when the premise holds—in the above example, ¬Wear. Each rule has associated (1) fitness indicating its suitability, (2) numerosity indicating the number of instances of the rule in the rule set, (3) predicted reward indicating the expected reward if the rule applies, and (4) prediction error.

XCS for Rationale Generation Briefly

The key features of XCS are *Rule discovery*, *Rules subsumption*, and *Action selection*. Rule discovery through the crossover and mutation processes involves introducing randomness to the antecedent by adding or removing factors, thereby generating rules that are either more general or specific. If a more general rule exists that exhibits lower predictive error within the given context, the algorithm will retain the more general rule and discard the more specific one. When selecting an action, the algorithm selects the one with the best-aggregated fitness. Supplemental material provides details of XCS.

An example of a rationale for not wearing a mask is {Risk=None, Preference=¬Wear, InteractWith=colleague}. This can be interpreted as mask is not needed when there is a no infection risk of and when agent prefers to not wear a mask while interacting with a colleague in the office. Each agent keeps the rules it discovers and evolves those in a rule set for decision-making.

Refining Rationale

Not all elements of a rule generated by XCS are appropriate to be a part of rationale. For instance, in Example 6, sharing personal preference is unnecessary when both agents value health. After generating the base rule, we refine the elements of rule considering the values of the actor and the observer agents. Following the example above, the agent who prefers the value of health then adjusts its rationale for the colleague who also cares about health to a health-related causal attribution, if it exists. For instance, no mask is required because of no risk of infection when interacting with a colleague in the office.

Algorithm 5 details the process of rationale generation. An agent first identifies rules associated with beliefs b_t (Line 2) and filters out rules not with the selected action (Line 3). The rationales are the aggregated rules (Line 4). An agent only reveals private information associated with the values of agents involved in the interaction (Line 5–7). For instance, if an agent who cares about freedom interacts with one who cares about freedom, it will exclude the infection risk from the environment in its rationales. For each rationale, an agent computes the privacy as the proportion of private attributes forming part of the rationale (Line 8).

5.3.4 Rationale Evaluation

On receiving a rationale from the actor agent, the receiver first updates its beliefs based on the rationale. Specifically, the receiver updates the beliefs of unobservable information from

Algorithm 5: Rationale generation

Input: beliefs b_t, Action a
Output: Rationale *Rat*
Function *GenRationale*:
    /* Generate associated rules with beliefs b_t */
    Get match set *ms* with b_t;
    Generate action set from *ms* with *a*;
    Aggregate rules *Rat* associated with action set;
    if *values not involved* **then**
        remove factors in *Rat* not related to presented values in b_t;
    end
    Compute privacy;
return

actor's context. In the rationale generation mask example, the receiver updates its beliefs of the infection risk to "None". When evaluating a rationale, the receiver makes an analogous decision based on the updated beliefs. If the receiver's computed action matches the actor's observed action in that context, the receiver accepts the actor's rationale.

Algorithm 6 details the evaluation of given rationales. Upon receiving a rationale, an agent reasons over the rationale. Specifically, the agent first updates its beliefs b_t based on the rationale in Line 3, precisely the private context or beliefs of others. With the provided rationale, an agent checks if any applicable rules align with its rule sets in Line 4. The agent identifies associated rules from b_t and adds them to applicable rules in Line 5. In Line 8, the agent calculates the fitness for each available action for each applicable rule and keeps the best action for each rule. The agent accepts this rationale if any selected action matches the observed action.

5.4 Simulation

We evaluate *Exanna* via a simulated pandemic scenario based on Examples 5 and 6 where agents move to various places, interact with other agents, decide to wear or not wear a mask, and provide a justification for their actions. We implemented our environment in MASON (Luke et al. 2005).

Algorithm 6: Evaluating a rationale

Input: Rationales *Rat*, Observed action a_{actor}, Beliefs b_t
Output: Decision d

Function *EvalRationale*:
    Initialize applicable rules *ars*;
    update b_t with private information in *Rat*;
    Add triggered match set from *Rat* to applicable rules *ars*;
    Add triggered match set from b_t to applicable rules *ars*;
    for *rule in applicable rules ars* **do**
        for *act in possible actions* **do**
            calculate fitness f_{act};
        end
        Keep the *act* with best fitness;
    end
    if *act contains* a_{actor} **then**
        Decision d = accept;
    else
        Decision d = reject;
    end
return

5.4.1 Scenario

The environment represents a multiagent society with several places and social circles. Our environment involves a finite population of 200 agents with different social circles. The environment has one park, one hospital, five homes, five offices, and five parties. Agents move around and interact in five places (home, office, party park, and hospital). Each agent is native to one home, one office, and one party. Agents in the same home, office, or party share the same family, colleague, or friend social circle. Each social circle has 40 agents. Time is represented in steps. Each agent moves to one place at each step and has a probability (50%) of interacting with one agent at the same place. Agents are more likely (75%) to move to places they are associated with when they move to home, office, and party, i.e., an agent is more likely to visit their own home than someone else's home.

Each agent forms its goal based on its value preferences. Specifically, each value in one context has a payoff matrix (Table 5.3 and 5.4); the weighted sum of the payoff determines the goal (desired states). An agent selecting an action not aligning with its goal, is considered deviating from its goal.

In the simulated environment, when an agent encounters another agent at the same place, it chooses an action based on its goal—whether to wear a mask. In addition, the agent justifies its behavior based on its beliefs in that context. For instance, the agent gives a rationale—{Risk=None, InteractWith=Colleague}—while not wearing a mask. The beliefs of an agent include public and private attributes. Each agent receives a payoff according to the interaction place for action selection as in Table 5.2(a). Wearing a mask at a hospital during a pandemic is desirable. Place and value preferences determine the payoff an agent gives to itself. An agent also gives sanctions as feedback to others based on their actions. The sanctions are based on the social circle. Table 5.2(b) lists the sanctions associated with social circles.

We run each simulation 10 times, and each simulation lasts 30,000 steps. We consider value of freedom and health. The value of freedom means agents would claim their *free will* and tend to follow their preferences.

Table 5.2: Payoff and feedback based on place and circle.

Table 5.2(a) Actor's payoff based on the place. Numbers reflect general expectations of places.

Places	Wear	¬Wear
Home	−0,25	0,25
Office	0,25	−0,25
Party	−0,25	0,25
Park	−0,5	0,5
Hospital	0,5	−0,5

Table 5.2(b) Feedback from an observer based on social circle.

Social Circle	Observer's response	
	Reject	Accept
Family	−1,00	1,00
Friend	−0,75	0,75
Coworker	−0,50	0,50
Stranger	−0,25	0,25

Table 5.3: Payoffs for the value of freedom depend on an agents' preferences.

Table 5.3(a) Payoffs corresponding to a preference for wearing a mask.

		Agent 2	
		Wear	¬ Wear
Agent 1	Wear	1,0	1,0
	¬ Wear	−1,0	−1,0

Table 5.3(b) Payoffs corresponding to a preference for not wearing a mask.

		Agent 2	
		Wear	¬ Wear
Agent 1	Wear	−1,0	−1,0
	¬ Wear	1,0	1,0

Table 5.4: Payoffs for the value of health. The numbers reflect how safe an agent feels.

		Infection risk	
		No risk	High risk
Action	Wear	0,0	1,0
	¬ Wear	0,0	−1,0

5.4.2 Contextual Properties

Whereas agents have limited observations on the environment, the context includes the place (home, office, party, park, and hospital) where interactions occur, the relationship (family, friend, colleague, and stranger) with the observer, the subjective belief of infection risk of the environment, the personal preference on mask-wearing, and the types of observer agents. Due to the partial observation, agents act based on their beliefs. Rationales enable belief updates.

5.4.3 Types of Societies

We define types of societies based on the rationale types. All societies include 50% of agents value health and 50% of agents value freedom. The value preferences of agents are as Table 5.5. All agents optimize their behavior based on the weighted sum of payoffs from themselves and others.

Baseline 1: Share All Society Agents share all information as rationales. Agents can make decision from the perspective of others.

Baseline 2: Share Decision Rules Society Agents share their decision rules as rationales.

***Exanna*: Share Value-Aligned Rules Society** Agents share their decision rules along with selective information that aligns with values as rationales.

Table 5.5: Value preferences of agents.

Agents: Values	Freedom	Health
Freedom-loving	1.0	0.0
Health-freak	0.0	1.0

5.4.4 Evaluation

We run simulations with share all, share decision rules, and *Exanna* societies. We propose the following hypotheses on resolution, social experience, privacy, and flexibility.

To test the hypotheses, we compute the following metrics.

$M_{\text{Resolution}}$	$\in [0, 100]$	Percentage of rationales accepted.
M_{Social}	$\in [-3, 3]$	Aggregate payoff that an agent receives for its behavior.
M_{Privacy}	$\in [0, 1]$	Proportion of private information retained during an interaction.
$M_{\text{Flexibility}}$	$\in [0, 1]$	Extent of deviation from an agent's own goal.

We conduct the independent t-test across the societies. We measure effect size with Glass's (1976) Δ since the societies have different standard deviations (Grissom and Kim 2012). We adopt Cohen's (1988) descriptors to interpret effect size: $<$0.2 indicates negligible, [0.2,0.5) indicates small, [0.5,0.8) indicate medium, and $>$0.8 indicates large effect.

5.5 Results

Table 5.6 summarizes the simulation results and the statistical analysis for our hypotheses. Summarily, *Exanna* offers better social experience, higher conflict resolution, and improved flexibility indicating that *Exanna* agents learn to act for greater societal good. The observed results follow our intuition that value-aligned rationales are more convincing. However, if an agent prefers to keep certain information private, deviation from goals is expected. Share All and Share Decision Rules societies results further indicate that more information in rationale is not always useful.

$H_{\text{Resolution}}$ Figure 5.1 compares conflict resolution in various societies. *Exanna* offers better conflict resolution ($p < 0.001$; $\Delta > 0.8$, indicating a large effect) than other societies. This result rejects the null hypothesis corresponding to $H_{\text{Resolution}}$. We observe that, in scenarios where providing a rationale do not convince others, *Exanna* agents are more flexible to deviate from their own goals to resolve conflicts.

$H_{\text{Social Experience}}$ For $H_{\text{Social Experience}}$, we measure the overall payoffs of agents in a society. An agent's payoff includes personal payoff from its action and the feedback from its interaction. Figure 5.2 compares the social experience for Share All, Share Decision Rules, and *Exanna* agent societies. We find that *Exanna* yields better social experience ($p < 0.001$; $\Delta > 0.8$, indicating a large effect)) than other societies. Specifically, *Exanna* agents receive better feedback from

Table 5.6: Results: Comparing mean ($\bar{X}$) and standard deviation (σ) of social experience, resolution, privacy, and flexibility in various societies. p is p-value from t-test.

		Share All	Share Rules	*Exanna*
$M_{Resolution}$	$\bar{X}$	0.58	0.58	0.60
	σ	0.02	0.01	0.02
	p	< 0.001	< 0.001	–
	Δ	1.80	3.11	–
M_{Social}	$\bar{X}$	0.59	0.62	0.70
	σ	0.06	0.02	0.05
	p	< 0.001	< 0.001	–
	Δ	1.80	3.11	–
$M_{Privacy}$	$\bar{X}$	0.00	1.34×10^{-5}	0.25
	σ	0.00	2.11×10^{-5}	2.90×10^{-3}
	p	< 0.001	< 0.001	–
	Δ	∞	11 896.52	–
$M_{Flexibility}$	$\bar{X}$	0.10	0.09	0.12
	σ	0.01	0.01	0.03
	p	0.08	< 0.01	–
	Δ	1.59	2.89	–

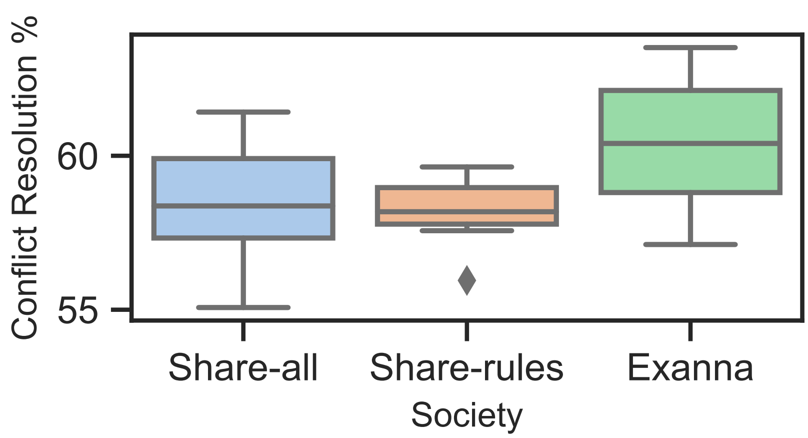


Figure 5.1: Comparing the resolution ($M_{Resolution}$) in various agent societies. The *Exanna* agent society has better resolution (Glass' $\Delta > 0.8$; $p < 0.001$) than the baseline societies.

other agents who receive their rationales. This result reject the null hypothesis that corresponds to $H_{Social\ Experience}$.

On closer analysis, we observe that *Exanna* agents receive more negative sanctions than other societies initially but soon learn to deviate from their goals.

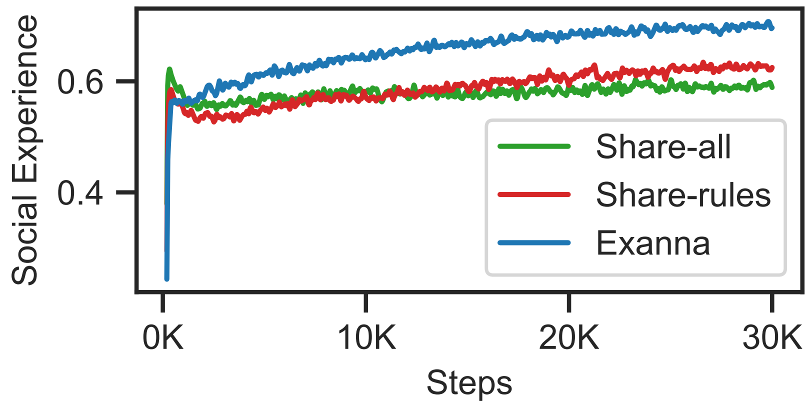


Figure 5.2: Comparing the social experience (M_{Social}) in various societies. *Exanna* agent society has better experience (Glass' $\Delta > 0.8$; $p < 0.001$) than other baselines.

$\mathbf{H_{Privacy}}$ *Exanna* agents better retain their privacy ($p < 0.001$; $\Delta > 0.8$, indicating a large effect) compared to Share All or Share Decision Rules agents. This analysis result reject the null hypothesis that corresponds to $H_{Privacy}$.

Although both Share Decision Rules and *Exanna* society share learned rules as rationales, *Exanna* agents aligns rationales to the receiver's values and limits the shared private information to values that agents appraise. A rationale stating causal attribution with minimum private information but aligned with agents' values is sufficient to explain behaviors.

$\mathbf{H_{Flexibility}}$ We compare agents' flexibility of goals to evaluate $H_{Flexibility}$. Figure 5.3 compares $M_{Flexibility}$ for Share All, Share Decision Rules, and *Exanna* agent societies. We find that *Exanna* offers higher flexibility ($p < 0.01$; $\Delta > 0.8$) than Share Decision Rules society. Although the mean flexibility in the Share All society is higher than in the *Exanna* society, this difference is not significant ($p > 0.05$).

Emerged Norm A norm emerges when the proportion of agents adhering to a particular behavior surpasses a threshold. We consider 90% as the threshold (Delgado 2002). We observe that compared to Share All and Share Decision Rules societies, *Exanna* promotes more general

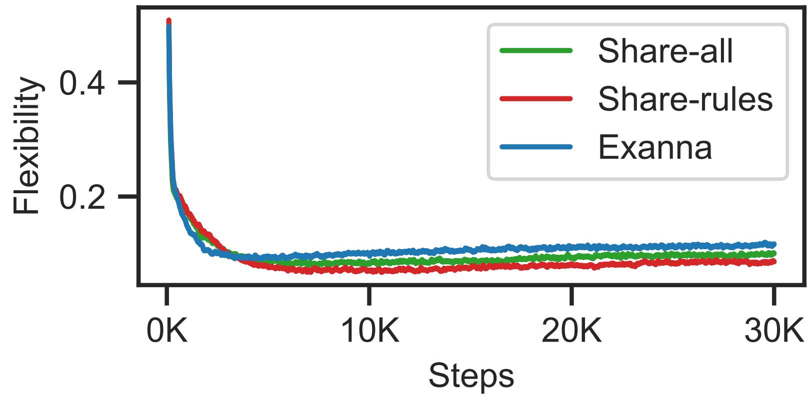


Figure 5.3: Comparing flexibility ($M_{\text{Flexibility}}$) in various agent societies. *Exanna* society shows higher flexibility (Glass' $\Delta > 0.8$; $p < 0.05$ for Share Decision Rules society but $p > 0.05$ for Share All society) than baseline societies.

norms. For instance, the following norms emerged only in *Exanna*.[1]

```
{preference = ¬Wear, InteractWith = Colleague,
location=OFFICE} => Wear
```

```
{ObererverAgentType = FREEDOM, InteractWith = Colleague,
location=HOSPITAL} => Wear
```

5.6 Conclusions and Directions

In human-centered AI systems, considerations for values are paramount in decision making and providing justifications for decisions made. AI agents must consider their stakeholders and accommodate human factors in their reasoning. We demonstrate via a multiagent study how we could create agents who incorporate values in decision making and in rationale generation and evaluation. Our results are consistent with our hypotheses. Value-aligned rationales offer better social experience and higher conflict resolution. While value-aligned rationales wrap partial information, agents learn to deviate from their goals to protect their privacy. Specifically, agents who receive rejections from others learn to be flexible to some extent to improve cooperation.

Assumptions and Limitations We make simplifying assumptions. First, agents can identify other agents' types. Second, agent types indicate their values which guide their behaviors. These

[1]The supplement includes the complete set of emerged norms.

assumptions may not apply in all cases but are essential when interacting with other agents.

While humans expect agent behaviors to align with human values, we assume value-aligned rationales best justify agent behaviors. In addition, although human values can change over time, we assume values will remain constant throughout our work.

In the simulation study, we limit value preferences to two values: health and freedom, to demonstrate how value preferences shape behaviors. A real-world scenario may include more intertwined values that are hard to quantify and tell which influences the decision more.

While the first step of value-aligned AI involves deducing values (Liscio et al. 2021, 2023) from stakeholders, our emphasis lies in illustrating how value-driven rationales shape agent behaviors.

Future Directions First, to include information cost in decision making. Different information may have different costs, which may change agents' final decisions. For instance, sharing tax data and sharing interest pose significantly different costs. An agent may be fine to share its interest but keep the tax data to itself. Second, to enable agents to decide what to share. In some cases, information suppression may be desirable. Third, to include rationales in decision-making instead of supplementary information. Having rationales as part of the decisions may increase the flexibility of an agent. Fourth, build an ontology to associate information with values, which we model as attributes. An ontology helps to model varied attributes or concepts and their intertwined relationships. While *Exanna* enables value-driven rationales and focuses on the decisions of a single agent, one future direction is to promote values and norms in a multiagent system.

CHAPTER

6

CONCLUSION

This dissertation tackles the challenges of accommodating humans in the loop, i.e., emotion as sanctions, social signals as responses to norms, and human values as guidance of behaviors. We present a framework for a dynamic MAS that actively involves human participation. Our framework aims to accommodate humans in the loop and operates in dynamic environments. The human factors we target are expressed emotions, social signals, social value orientation, and values.

6.1 Answering the Research Questions

Chapter 2 presents *Noe*, an agent framework that comprises emotional responses to the normative reasoning process. We show how emotion modeling in *Noe* enables the promotion of norm compliance and the improvement of societal welfare.

In Chapter 3, we define social signals as *sanction*, *tell*, and *hint*. *Ness* gives an agent framework that models normative information from social signals to support norm emergence. Modeling soft signals such as hints or messages prevents unfavorable outcomes, such as negative sanctions and straying from goals, while leading to greater satisfaction than the baseline agent societies despite requiring an equivalent amount of information.

Fleur (Chapter 4) presents a framework that operationalizes the concept of Social Value Orientation (SVO). SVO provides agents with different preferences over resource allocations between themselves and others Aligning with SVO enables better social experience and robust norm emergence.

Chapter 5 proposes *Exanna*, a framework that incorporates values in decision-making, rationale generation, and reasoning over rationale. Constructing rationales based on agent values enhances social experience and conflict resolution with some tradeoffs with goal deviation.

6.2 Future Directions

This work suggests numerous significant and captivating expansions. This section delineates two primary forthcoming directions related to this dissertation.

Human-Centered Autonomous Systems

This research delves into investigating the impact of various elements of human factors (including emotional responses, social signals, social value orientation, and human values) on the decision-making process of an individual agent. Our future directions include further understanding the causal connections between decisions and human factors.

An intriguing direction involves exploring how social signals and various social norms are interconnected. Furthermore, while our work focuses on individual agent behaviors, expanding agent modeling from the micro to the macro level can provide insights into comprehending and building pertinent norms. To be more precise, the micro level refers to an individual agent's perspective, while the macro level pertains to perspectives within multiagent systems (Woodgate and Ajmeri 2022). One possible beginning for expanding upon this study involves investigating the process of modifying norms to foster particular values within MAS.

Evolution of Social Norms

Intuitively, individuals driven by self-interest would not take actions to achieve the collective interest. However, real-world observations often defy this intuition. This gap has piqued the interest of researchers in the field of social science. Since Axelrod and Hamilton (1981) propose the model of the evolution of cooperation, evolutionary theory has been applied to study the evolution of cooperation (Axelrod and Dion 1988; Nowak 2006) or collective actions (social norms) (Poteete et al. 2010). By combining principles from evolutionary biology and game theory, evolutionary game theory (EGT) studies the dynamics of social behaviors among a population within the framework of evolutionary processes over time (Hofbauer and Sigmund 1998). An intriguing direction is studying the emergence and survival of agents with different values.

Rationales Behind the Outcomes of AI Systems

Our future research extensions include delving deeper into the costs associated with information and deliberate information concealment to provide rationales for decisions. In reality, the value of information can vary significantly. Exploring the impact of varying information costs on decisions could enhance the accuracy and reliability of action recommendations and the formulation of rationales. In terms of information concealment, information suppression may be permissible for the sake of strategic advantage, security and, ethical considerations. Formulating flexible rationales based on contextual property is one promising direction. Lastly, our future directions include developing agents capable of determining both the content and timing of information sharing, thereby enhancing the flexibility of strategies.

REFERENCES

Rishabh Agrawal, Nirav Ajmeri, and Munindar P. Singh. Socially intelligent genetic agents for the emergence of explicit norms. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 10–14, Vienna, July 2022. IJCAI. doi: 10.24963/ijcai.2022/2.

Stéphane Airiau, Sandip Sen, and Daniel Villatoro. Emergence of conventions through social learning. *Autonomous Agents and Multi-Agent Systems (JAAMAS)*, 28(5):779–804, 2014. doi: 10.1007/s10458-013-9237-x.

Nirav Ajmeri, Pradeep K. Murukannaiah, Hui Guo, and Munindar P. Singh. Arnor: Modeling social intelligence via norms to engineer privacy-aware personal agents. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 230–238, São Paulo, May 2017. IFAAMAS. doi: 10.5555/3091125.3091163.

Nirav Ajmeri, Hui Guo, Pradeep K. Murukannaiah, and Munindar P. Singh. Robust norm emergence by revealing and reasoning about context: Socially intelligent agents for enhancing privacy. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 28–34, Stockholm, July 2018. IJCAI. doi: 10.24963/ijcai.2018/4.

Nirav Ajmeri, Hui Guo, Pradeep K. Murukannaiah, and Munindar P. Singh. Elessar: Ethics in norm-aware agents. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems, (AAMAS)*, pages 16–24, Auckland, May 2020. IFAAMAS. doi: 10.5555/3398761.3398769.

Bexy Alfonso Espinosa. *Agents with Affective Traits for Decision-Making in Complex Environments*. PhD thesis, Universitat Politècnica de València, 2017.

Giulia Andrighetto, Jordi Brandts, Rosaria Conte, Jordi Sabater-Mir, Hector Solaz, and Daniel Villatoro. Punish and voice: Punishment enhances cooperation when combined with norm-signalling. *PLOS ONE*, 8(6):1–8, 2013. doi: 10.1371/journal.pone.0064941.

Suwardi Annas, Muh Isbar Pratama, Muh Rifandi, Wahidah Sanusi, and Syafruddin Side. Stability analysis and numerical simulation of SEIR model for pandemic COVID-19 spread in Indonesia. *Chaos, Solitons & Fractals*, 139:110072, 2020. doi: 10.1016/j.chaos.2020.110072.

Estefania Argente, Elena Del Val, Daniel Perez-Garcia, and Vicente Botti. Normative emotional agents: A viewpoint paper. *IEEE Transactions on Affective Computing*, 13(3):1254–1273, 2022. doi: 10.1109/TAFFC.2020.3028512.

Robert Axelrod and Douglas Dion. The further evolution of cooperation. *Science*, 242(4884): 1385–1390, 1988. doi: 10.1126/science.242.4884.138.

Robert Axelrod and William D. Hamilton. The evolution of cooperation. *Science*, 211(4489): 1390–1396, 1981. doi: 10.1126/science.74663.

Gonul Ayci, Pınar Yolum, Arzucan Özgür, and Murat Şensoy. Explain to me: Towards understanding privacy decisions. In *Proceedings of the 22th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 2790—-2791, London, 2023. IFAAMAS. doi: 10.5555/3545946.3599079.

Lisa Feldman Barrett, Ralph Adolphs, Stacy Marsella, Aleix M. Martinez, and Seth D. Pollak. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1):1–68, 2019.

Mathieu Bourgais, Patrick Taillandier, and Laurent Vercouter. Ben: An agent architecture for explainable and expressive behavior in social simulation. In *International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems*, pages 147–163, Cham, 2019. Springer. doi: 10.1007/978-3-030-30391-4_9.

Jan Broersen, Mehdi Dastani, Joris Hulstijn, Zisheng Huang, and Leendert van der Torre. The BOID architecture: Conflicts between beliefs, obligations, intentions and desires. In *Proceedings of the 5th International Conference on Autonomous Agents*, pages 9–16, New York, NY, USA, 2001. Association for Computing Machinery. doi: 10.1145/375735.375766.

Martin V. Butz and Stewart W. Wilson. An algorithmic description of XCS. In *Proceedings of the 3rd International Workshop on Learning Classifier Systems*, volume 1996 of *LNCS*, pages 253–272, Paris, France, 2000. Springer. doi: 10.1007/3-540-44640-0_15.

Gary Charness and Matthew Rabin. Understanding social preferences with simple tests. *The quarterly journal of economics*, 117(3):817–869, 2002.

Amit K. Chopra and Munindar P. Singh. From social machines to social protocols: Software engineering foundations for sociotechnical systems. In *Proceedings of the 25th International*

World Wide Web Conference, pages 903–914, Montréal, April 2016. ACM. doi: 10.1145/2872427.2883018.

Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 2nd edition, 1988. doi: 10.4324/9780203771587.

Victor Contreras, Michael Schumacher, and Davide Calvaresi. Integration of local and global features explanation with global rules extraction and generation tools. In *Explainable and Transparent AI and Multi-Agent Systems (EXTRAAMAS)*, volume 13283 of *LNCS*, pages 19–37, Virtual Conference, 2022. Springer. doi: 10.1007/978-3-031-15565-9_2.

Igor Conrado Alves de Lima, Luis Gustavo Nardin, and Jaime Simão Sichman. Gavel: A sanctioning enforcement framework. In *Proceedings of the 6th International Workshop on Engineering Multi-Agent Systems (EMAS)*, pages 225–241, Cham, 2019. Springer. doi: 10.1007/978-3-030-25693-7_12.

Jan de Mooij, Davide Dell'Anna, Parantapa Bhattacharya, Mehdi Dastani, Brian Logan, and Samarth Swarup. Quantifying the effects of norms on COVID-19 cases using an agent-based simulation. In *Proceedings of the The 22nd International Workshop on Multi-Agent-Based Simulation (MABS)*, pages 99–112, Cham, 2022. Springer International Publishing. doi: 10.1007/978-3-030-94548-0_8.

Carolyn H. Declerck and Sandy Bogaert. Social value orientation: Related to empathy and the ability to read the mind in the eyes. *The Journal of Social Psychology*, 148(6):711–726, 2008. doi: 10.3200/SOCP.148.6.711-726.

Jordi Delgado. Emergence of social conventions in complex networks. *Artificial Intelligence*, 141(1–2):171–185, 2002. doi: 10.1016/S0004-3702(02)00262-X.

Davide Dell'Anna, Mehdi Dastani, and Fabiano Dalpiaz. Runtime revision of norms and sanctions based on agent preferences. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1609–1617. IFAAMAS, 2019. doi: 10.5555/3306127.3331881.

Davide Dell'Anna, Mehdi Dastani, and Fabiano Dalpiaz. Runtime revision of sanctions in normative multi-agent systems. *Autonomous Agents and Multi-Agent Systems (JAAMAS)*, 34 (2):43:1–43:54, June 2020. doi: 10.1007/s10458-020-09465-8.

Daan Di Scala and Pınar Yolum. Paccart: Reinforcing trust in multiuser privacy agreement systems. In *Proceedings of the 22th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 2787–2789, London, 2023. IFAAMAS. doi: 10.5555/3545946.3599078.

Frank Dignum, Virginia Dignum, Paul Davidsson, Amineh Ghorbani, Mijke van der Hurk, Maarten Jensen, Christian Kammler, Fabian Lorig, Luis Gustavo Ludescher, Alexander Melchior, René Mellema, Cezara Pastrav, Loïs Vanhee, and Harko Verhagen. Analysing the combined health, social and economic impacts of the corovanvirus pandemic using agent-based social simulation. *Minds and Machines*, 30(2):177–194, 2020. doi: 10.1007/s11023-020-09527-6.

Ward Edwards. The theory of decision making. *Psychological Bulletin*, 51(4):380, 1954. doi: 10.1037/h0053870.

Christopher Frantz and Gabriella Pigozzi. Modeling norm dynamics in multiagent systems. *Journal of Applied Logics – IfCoLoG Journal of Logics and their Applications*, 5(2):491–564, 2018.

Athina Georgara, Juan A. Rodriguez Aguilar, and Carles Sierra. Building contrastive explanations for multi-agent team formation. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 516–524, Auckland, 2022. IFAAMAS. doi: 10.5555/3535850.3535909.

Adam Gerace. Internal and external attributions. In Virgil Zeigler-Hill and Todd K. Shackelford, editors, *Encyclopedia of Personality and Individual Differences*, pages 2328–2334. Springer, 2020. doi: 10.1007/978-3-319-24612-3_2301.

Gene V. Glass. Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5 (10):3–8, 1976. doi: 10.3102/0013189X005010003.

Donald W. Griesinger and James W. Livingston Jr. Toward a model of interpersonal motivation in experimental games. *Behavioral Science*, 18(3):173–188, 1973. doi: 10.1002/bs.3830180305.

Robert J. Grissom and John J. Kim. *Effect Sizes for Research: Univariate and Multivariate Applications*. Routledge, Abingdon-on-Thames, 2012. doi: 10.4324/9780203803233.

Jayesh K. Gupta, Maxim Egorov, and Mykel Kochenderfer. Cooperative multi-agent control using deep reinforcement learning. In *Proceedings of the 9th Workshop on Adaptive Learning*

Agents workshop, volume 10642 of *LNCS*, pages 66–83, São Paulo, 2017. Springer. doi: 10.1007/978-3-319-71682-4_5.

Jianye Hao, Jun Sun, Guangyong Chen, Zan Wang, Chao Yu, and Zhong Ming. Efficient and robust emergence of norms through heuristic collective learning. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, 12(4):1–20, 2017. doi: 10.1145/3127498.

Michael Hind, Dennis Wei, Murray Campbell, Noel C.F. Codella, Amit Dhurandhar, Aleksandra Mojsilović, Karthikeyan Natesan Ramamurthy, and Kush R. Varshney. TED: Teaching AI to explain its decisions. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 123–129, Honolulu, 2019. ACM. doi: 10.1145/3306618.3314273.

Josef Hofbauer and Karl Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998. doi: 10.1017/CBO9781139173179.

Christopher D. Hollander and Annie S. Wu. The current state of normative agent-based systems. *Journal of Artificial Societies and Social Simulation*, 14(2):6, 2011. doi: 10.18564/jasss.1750.

Özgür Kafalı, Nirav Ajmeri, and Munindar P. Singh. Revani: Revising and verifying normative specifications for privacy. *IEEE Intelligent Systems (IS)*, 31(5):8–15, September 2016. doi: 10.1109/MIS.2016.89.

Özgür Kafalı, Nirav Ajmeri, and Munindar P. Singh. DESEN: Specification of sociotechnical systems via patterns of regulation and control. *TOSEM*, 29(1):7:1–7:50, February 2020. doi: 10.1145/3365664.

Anup K. Kalia, Nirav Ajmeri, Kevin Chan, Jin-Hee Cho, Sibel Adalı, and Munindar P. Singh. The interplay of emotions and norms in multiagent systems. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 371–377, Macau, August 2019. IJCAI. doi: 10.24963/ijcai.2019/53.

Dacher Keltner and Jonathan Haidt. Social functions of emotions at four levels of analysis. *Cognition and Emotion*, 13(5):505–521, 1999. doi: 10.1080/026999399379168.

A. Kurtan and Pınar Yolum. Assisting humans in privacy management: An agent-based approach. *Autonomous Agents and Multi-Agent Systems*, 35(1):1–33, 2021. doi: 10.1007/s10458-020-09488-1.

Pat Langley. Explainable, normative, and justified agency. In *Proceedings of the 33nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 9775–9779, Honolulu, Hawaii, USA, 2019. AAAI Press. doi: 10.1609/aaai.v33i01.33019775.

Roger Lera-Leri, Filippo Bistaffa, Marc Serramia, Maite Lopez-Sanchez, and Juan Rodriguez-Aguilar. Towards pluralistic value alignment: Aggregating value systems through ℓp-regression. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 780–788, Auckland, 2022. IFAAMAS. doi: 10.5555/3535850.3535938.

Priel Levy and Nathan Griffiths. Convention emergence with congested resources. In *Proceedings of the 19th European Conference on Multi-Agent Systems*, pages 126–143, Cham, 2021. Springer. doi: 10.1007/978-3-030-82254-5_8.

Beishui Liao, Pere Pardo, Marija Slavkovik, and Leendert van der Torre. The jiminy advisor: Moral agreements among stakeholders based on norms and argumentation. *Journal of Artificial Intelligence Research (JAIR)*, 77:737–792, 2023. doi: 10.1613/jair.1.14368.

Enrico Liscio, Michiel van der Meer, Luciano C. Siebert, Catholijn M. Jonker, Niek Mouter, and Pradeep K. Murukannaiah. Axies: Identifying and evaluating context-specific values. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 799–808, London, 2021. IFAAMAS. doi: 10.5555/3463952.3464048.

Enrico Liscio, Roger Lera-Leri, Filippo Bistaffa, Roel I.J. Dobbe, Catholijn M. Jonker, Maite Lopez-Sanchez, Juan A. Rodriguez-Aguilar, and Pradeep K. Murukannaiah. Value inference in sociotechnical systems. In *Proceedings of the 22th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1774–1780, London, 2023. IFAAMAS. doi: 10.5555/3545946.3598838.

Maite Lopez-Sanchez, Marc Serramia, Juan A Rodriguez-Aguilar, Javier Morales, and Michael Wooldridge. Automating decision making to help establish norm-based regulations. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1613–1615, São Paulo, May 2017. IFAAMAS. doi: 10.5555/3091125.3091380.

Sean Luke, Claudio Cioffi-Revilla, Liviu Panait, Keith Sullivan, and Gabriel Balan. MASON: A multiagent simulation environment. *Simulation: Transactions of the Society for Modeling and Simulation International*, 81(7):517–527, July 2005.

Javier Marín-Morales, Juan Luis Higuera-Trujillo, Alberto Greco, Jaime Guixeres, Carmen Llinares, Enzo Pasquale Scilingo, Mariano Alcañiz, and Gaetano Valenza. Affective computing in virtual reality: Emotion recognition from brain and heartbeat dynamics using wearable sensors. *Scientific Reports*, 8(1):1–15, 2018. doi: 10.1038/s41598-018-32063-4.

Ofir Marom and Benjamin Rosman. Belief reward shaping in reinforcement learning. In *Proceedings of the 32nd Conference on Artificial Intelligence (AAAI)*, pages 3762–3769, New Orleans, 2018. AAAI Press. doi: 10.5555/3504035.3504496.

Stacy Marsella, Jonathan Gratch, and Paolo Petta. Computational models of emotion. In Klaus R. Scherer, Tanja Banziger, and Etienne Roesch, editors, *Blueprint for Affective Computing*, chapter 1.2, pages 21–41. Oxford University Press, 2010.

Stacy C. Marsella and Jonathan Gratch. EMA: A process model of appraisal dynamics. *Cognitive Systems Research*, 10(1):70–90, 2009. doi: 10.1016/j.cogsys.2008.03.005.

David Masad and Jacqueline Kazil. MESA: An agent-based modeling framework. In *Proceedings of the 14th PYTHON in Science Conference*, pages 53–60, Austin, 2015. Springer.

Mehdi Mashayekhi, Nirav Ajmeri, George F. List, and Munindar P. Singh. Prosocial norm emergence in multiagent systems. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, 17(1–2):3:1–3:24, June 2022. doi: 10.1145/3540202.

Kevin R. McKee, Ian M. Gemp, Brian McWilliams, Edgar A. Duéñez-Guzmán, Edward Hughes, and Joel Z. Leibo. Social diversity and social preferences in mixed-motive reinforcement learning. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 869–877, Auckland, 2020. IFAAMAS. doi: 10.5555/3398761.3398863.

Stanley Milgram, Hilary J. Liberty, Raymond Toledo, and Joyce Wackenhut. Response to intrusion into waiting lines. *Journal of Personality and Social Psychology*, 51(4):683, 1986. doi: 10.1037/0022-3514.51.4.683.

Tim Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267:1–38, 2019. doi: 10.1016/j.artint.2018.07.007.

Thomas M. Moerland, Joost Broekens, and Catholijn M. Jonker. Emotion in reinforcement learning agents and robots: A survey. *Machine Learning*, 107(2):443–480, 2018. doi: 10.1007/s10994-017-5666-0.

Nieves Montes and Carles Sierra. Value-guided synthesis of parametric normative systems. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 907–915, London, 2021. IFAAMAS. doi: 10.5555/3463952.3464060.

Nieves Montes and Carles Sierra. Synthesis and properties of optimally value-aligned normative systems. *Journal of Artificial Intelligence Research (JAIR)*, 74:1739–1774, 2022. doi: 10.1613/jair.1.13487.

Javier Morales, Michael J. Wooldridge, Juan A. Rodríguez-Aguilar, and Maite López-Sánchez. Off-line synthesis of evolutionarily stable normative systems. *Autonomous Agents and Multi-Agent Systems (JAAMAS)*, 32(5):635–671, 2018. doi: 10.1007/s10458-018-9390-3.

Andreasa Morris-Martin, Marina De Vos, and Julian Padget. Norm emergence in multiagent systems: A viewpoint paper. *Autonomous Agents and Multi-Agent Systems (JAAMAS)*, 33(6): 706–749, 2019. doi: 10.1007/s10458-019-09422-0.

Francesca Mosca and Jose M. Such. ELVIRA: An explainable agent for value and utility-driven multiuser privacy. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 916–924, London, May 2021. IFAAMAS. doi: 10.5555/3463952.3464061.

Ryan O. Murphy and Kurt A. Ackermann. Social value orientation: Theoretical and measurement issues in the study of social preferences. *Personality and Social Psychology Review*, 18(1): 13–41, 2014. doi: 10.1177/1088868313501745.

Pradeep K. Murukannaiah, Nirav Ajmeri, Catholijn M. Jonker, and Munindar P. Singh. New foundations of ethical multiagent systems. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1706–1710, Auckland, May 2020. IFAAMAS. doi: 10.5555/3398761.3398958. Blue Sky Ideas Track.

Luis G. Nardin, Tina Balke-Visser, Nirav Ajmeri, Anup K. Kalia, Jaime S. Sichman, and Munindar P. Singh. Classifying sanctions and designing a conceptual sanctioning process model for socio-technical systems. *The Knowledge Engineering Review (KER)*, 31(2):142–166, March 2016. doi: 10.1017/S0269888916000023.

Martin A. Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560–1563, 2006. doi: 10.1126/science.11337.

Gideon Ogunniye and Nadin Kökciyan. Contextual integrity for argumentation-based privacy reasoning. In *Proceedings of the 22th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 2253–2261, London, 2023. IFAAMAS. doi: 10.5555/3545946.3598903.

Andrew Ortony, Gerald L. Clore, and Allan Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, New York, 1988. doi: 10.1017/CBO9780511571299.

Michael Pernpeintner. Toward a self-learning governance loop for competitive multi-attribute MAS. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems AAMAS*, pages 1619–1621, London, 2021. IFAAMAS. doi: 10.5555/3463952.3464179.

Piero Poletti, Marcello Tirani, Danilo Cereda, Filippo Trentini, Giorgio Guzzetta, Giuliana Sabatino, Valentina Marziano, Ambra Castrofino, Francesca Grosso, Gabriele Del Castillo, Raffaella Piccarreta, ATS Lombardy COVID-19 Task Force, Aida Andreassi, Alessia Melegaro, Maria Gramegna, Marco Ajelli, and Stefano Merler. Probability of symptoms and critical disease after SARS-CoV-2 infection. *arXiv*, 2006.08471:1–10, 2020. URL `https://arxiv.org/abs/2006.08471`.

Amy R. Poteete, Marco A. Janssen, and Elinor Ostrom. *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*. Princeton University Press, 2010. doi: 10.1515/9781400835157.

Anand S. Rao and Michael P. Georgeff. Modeling rational agents within a BDI-architecture. In *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 473–484, San Francisco, 1991. Morgan Kaufmann. doi: 10.5555/3087158.3087205.

John Realpe-Gómez, Giulia Andrighetto, Luis Gustavo Nardin, and Javier Antonio Montoya. Balancing selfishness and norm conformity can explain human behavior in large-scale prisoner's dilemma games and can poise human groups near criticality. *Physical Review E.*, 97 (4):042321, 2018. doi: 10.1103/PhysRevE.97.042321.

Rudolph J. Rummel. Understanding conflict and war: Vol. 1: The dynamic psychological field. *Sage Publications*, 1975.

Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson, New Jersey, 3 edition, 2010.

Bastin Tony Roy Savarimuthu and Stephen Cranefield. Norm creation, spreading and emergence: A survey of simulation models of norms in multi-agent systems. *Multiagent and Grid Systems*, 7(1):21–54, 2011. doi: 10.3233/MGS-2011-0167.

Bastin Tony Roy Savarimuthu, Stephen Cranefield, Maryam A. Purvis, and Martin K. Purvis. Obligation norm identification in agent societies. *Journal of Artificial Societies and Social Simulation*, 13(4):3, 2010. doi: 10.18564/jasss.1659.

Shalom H. Schwartz. An overview of the schwartz theory of basic values. *Online readings in Psychology and Culture*, 2(1):2307–0919, 2012. doi: 10.9707/2307-0919.1116.

Norbert Schwarz. Emotion, cognition, and decision making. *Cognition and Emotion*, 14(4): 433–440, 2000. doi: 10.1080/026999300402745.

Marc Serramia, Maite López-Sánchez, Juan A. Rodríguez-Aguilar, Manel Rodríguez, Michael Wooldridge, Javier Morales, and Carlos Ansótegui. Moral values in norm decision making. In *Proceedings of the 17th Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1294–1302, Stockholm, July 2018. IFAAMAS. doi: 10.5555/3237383.3237891.

Herbert A. Simon. *The New Science of Management Decision*. Harper & Brothers, 1960. doi: 10.1037/13978-000.

Herbert A. Simon. Motivational and emotional controls of cognition. *Psychological Review*, 74 (1):29–39, 1967. doi: 10.1037/h0024127.

Amika M. Singh and Munindar P. Singh. Norm deviation in multiagent systems: A foundation for responsible autonomy. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1–9, Macau, August 2023. IJCAI.

Munindar P. Singh. *Multiagent Systems: A Theoretical Framework for Intentions, Know-How, and Communications*. Number 799 in LNCS. Springer, Heidelberg, 1994. doi: 10.1007/BFb0030531. URL `http://www.csc.ncsu.edu/faculty/mpsingh/books/MAS/`.

Munindar P. Singh. Norms as a basis for governing sociotechnical systems. *TIST*, 5(1): 21:1–21:23, December 2013. doi: 10.1145/2542182.2542203.

Paul Slovic. The construction of preference. *American Psychologist*, 50(5):364, 1995. doi: 10.1037/0003-066X.50.5.364.

Aron Szekely, Francesca Lipari, Alberto Antonioni, Mario Paolucci, Angel Sánchez, Luca Tummolini, and Giulia Andrighetto. Evidence from a long-term experiment that collective risks change social norms and promote cooperation. *Nature Communications*, 12:5452:1–5452:7, September 2021. doi: 10.1038/s41467-021-25734-w.

Myrthe L. Tielman, Catholijn M. Jonker, and M. Birna Van Riemsdijk. Deriving norms from actions, values and context. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, pages 2223–2225, 2019. doi: 10.5555/3306127.3332065.

Sz-Ting Tzeng, Nirav Ajmeri, and Munindar P. Singh. Noe: Norms emergence and robustness based on emotions in multiagent systems. In *Proceedings of the International Workshop on Coordination, Organizations, Institutions, Norms and Ethics for Governance of Multi-Agent Systems (COINE)*, volume 13239 of *LNCS*, pages 62–77, Virtual, May 2021. Springer. doi: 10.1007/978-3-031-16617-4_5.

Ryan J. Urbanowicz and Will N. Browne. *Introduction to Learning Classifier Systems*. Springer Briefs in Intelligent Systems. Springer, New York, 2017. doi: 10.1007/978-3-662-55007-6.

Paul A. M. Van Lange. The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology*, 77(2): 337–349, August 1999. doi: 10.1037/0022-3514.77.2.337.

Francisco J. Varela, Evan Thompson, and Eleanor Rosch. *The Embodied Mind: Cognitive Science and Human Experience*. MIT press, 2017. doi: 10.7551/mitpress/6730.001.0001.

Henricus J.E. Verhagen. *Norm autonomous agents*. PhD thesis, Stockholm Universitet, 2000.

Christian von Scheve, Daniel Moldt, Julia Fix, and Rolf von Luede. My agents love to conform: Norms and emotion in the micro-macro link. *Computational and Mathematical Organization Theory*, 12(2–3):81–100, 2006. doi: 10.1007/s10588-006-9538-6.

Georg Henrik Von Wright. *Norm and Action: A Logical Enquiry*. International Library of Philosophy and Scientific Method. Humanities Press, New York, 1963.

Eric Wang, Pasha Khosravi, and Guy Van den Broeck. Probabilistic sufficient explanations. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 3082–3088, Montreal, August 2021. IJCAI. doi: 10.24963/ijcai.2021/424.

Christopher J.C.H. Watkins and Peter Dayan. Q-learning. *Machine Learning*, 8(3–4):279–292, 1992. doi: 10.1007/BF00992698.

Michael Winikoff, Galina Sidorenko, Virginia Dignum, and Frank Dignum. Why bad coffee? explaining BDI agent behaviour with valuings. *Artificial Intelligence*, 300:103554, 2021. doi: 10.1016/j.artint.2021.103554.

Jessica Woodgate and Nirav Ajmeri. Macro ethics for governing equitable sociotechnical systems. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1824–1828, Online, May 2022. IFAAMAS. doi: 10.5555/3535850.3536118. Blue Sky Ideas Track.

Yang Wu and Laura E Schulz. Understanding social display rules: Using one person's emotional expressions to infer the desires of another. *Child Development*, 91(5):1786–1799, 2020. doi: 10.1111/cdev.13346.

Yang Wu, Chris L. Baker, Joshua B. Tenenbaum, and Laura E Schulz. Rational inference of beliefs and desires from emotional expressions. *Cognitive Science*, 42(3):850–884, 2018. doi: 10.1111/cogs.12548.

Chayu Yang and Jin Wang. A mathematical model for the novel coronavirus epidemic in Wuhan, China. *Mathematical Biosciences and Engineering MBE*, 17(3):2708, 2020. doi: 10.3934/mbe.2020148.

Vahid Yazdanpanah, Enrico H. Gerding, Sebastian Stein, Mehdi Dastani, Catholijn M. Jonker, and Timothy J. Norman. Reasoning about responsibility in autonomous systems: Challenges and opportunities. *AI & Society: Journal of Knowledge, Culture and Communication*, 2023. doi: 0.1007/s00146-022-01607-8. In press.

APPENDICES

.1 Appendix: *Ness*

.1.1 Reproducibility

Table 1 lists the parameters in our simulation.

Table 1: Hyperparameters.

Parameter	Value	Comment
Learning rate α	0.001	
Discount factor γ	0.9	
Simulation step per action	1	
Population size	100	
Infection %	0.3	The default fraction of infected agents in a society
Certainty of potential reward	0.3	Value for κ for certainty of possible sanctions from normative information through hints
Certainty of potential reward	0.5	Value for κ for certainty of possible sanctions from normative information messages

.1.2 Additional Results

Figure 1 plots the total number of infected agents in various societies.

Figures 2, 3, 4, and 5 shows plots for the number of infected, deceased, healthy, and vaccinated agents in the first 500 steps in various societies, where the differences are noticeable.

Figure 6 shows the number of agents in isolation and quarantine in various societies in the first 500 steps of the simulation.

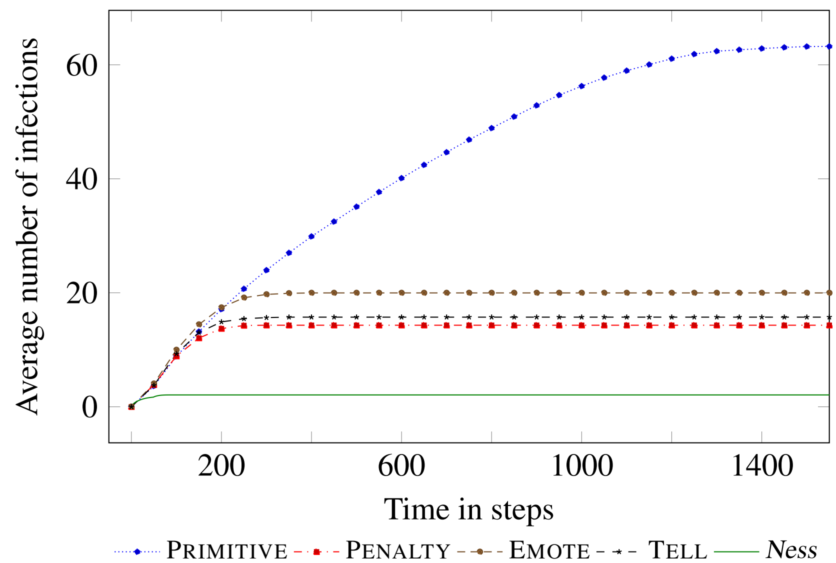


Figure 1: Comparing the average number of infections ($M_{\text{Infections}}$) in various societies. *Ness* yields fewer infections on average than other societies. The effect is large.

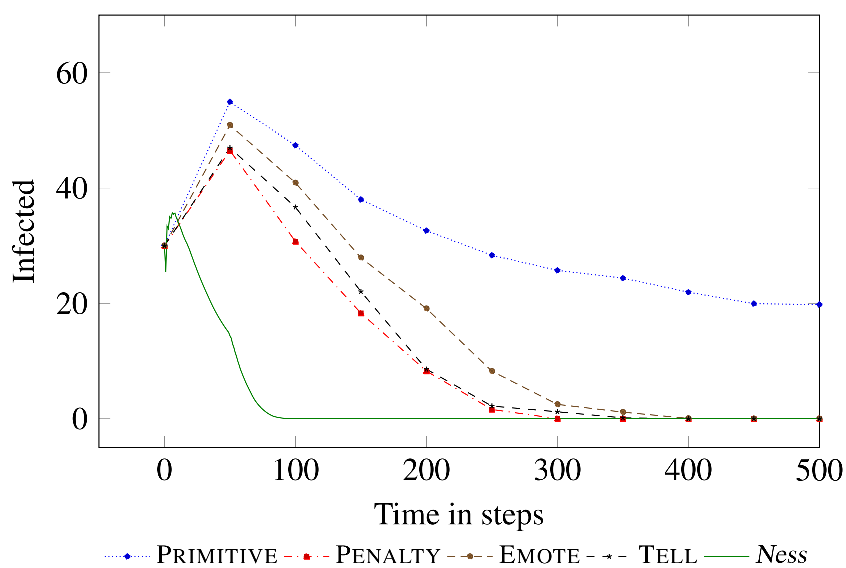


Figure 2: Comparing the number of infected agents in the first 500 steps. The differences are noticeable here.

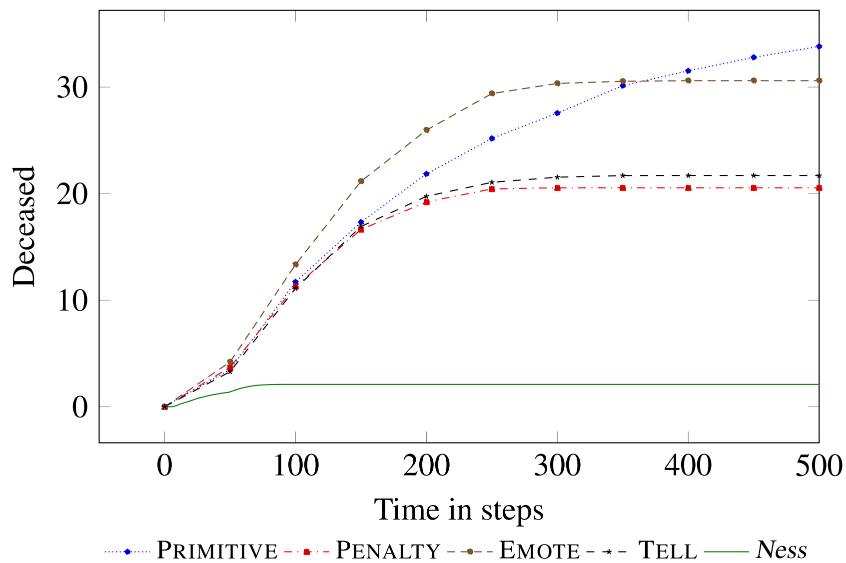


Figure 3: Comparing the number of deceased agents in the first 500 steps. The differences between *Ness* and EMOTE are noticeable here.

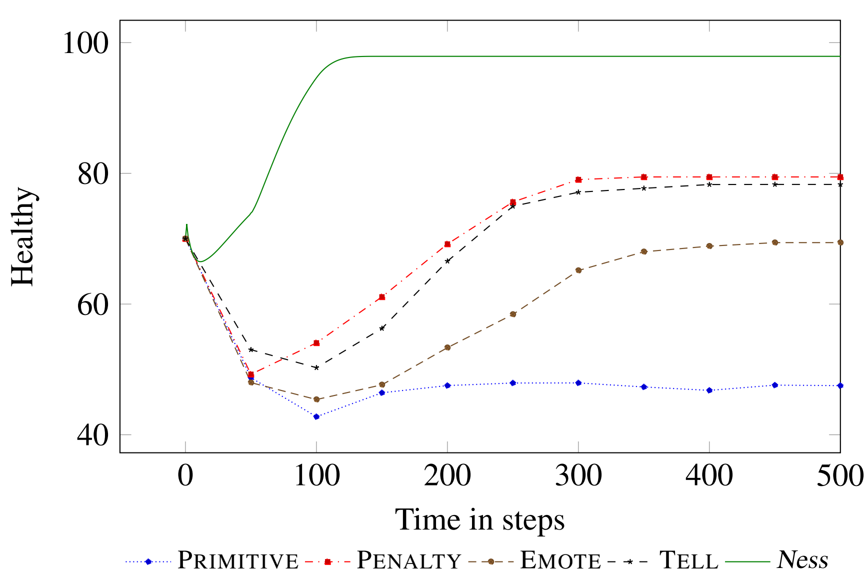


Figure 4: Comparing the number of healthy agents in the first 500 steps. The differences are noticeable here.

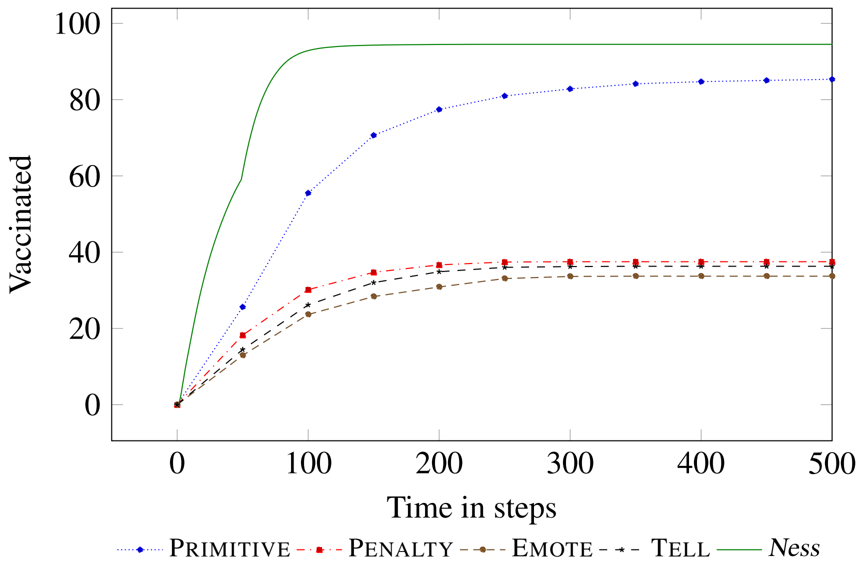


Figure 5: Comparing the number of vaccinated agents in the first 500 steps. The differences between *Ness* and other societies are noticeable here.

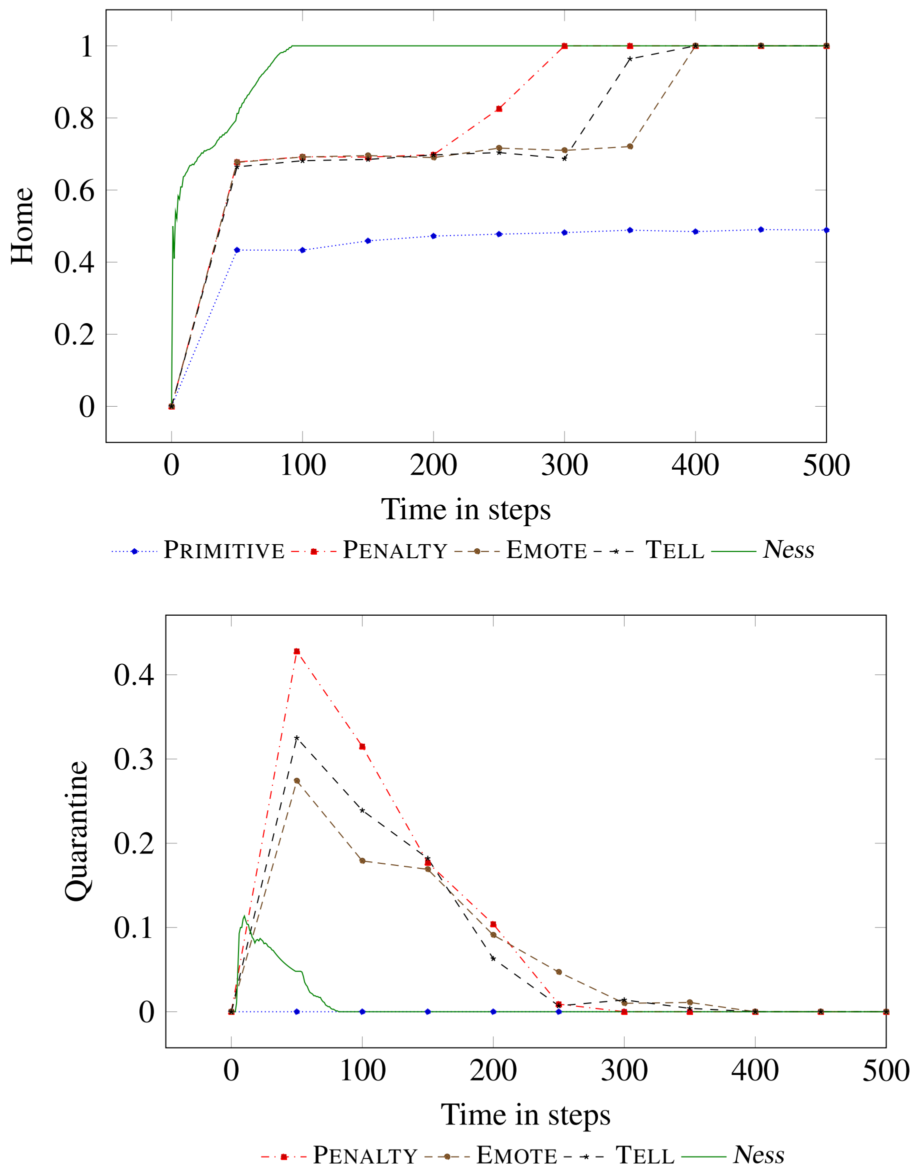


Figure 6: Comparing the number of agents in isolation (M_{Home}) and the number of agents in quarantine ($M_{Quarantine}$) in various societies in the first 500 steps.

.2 Appendix: *Exanna*

.2.1 Appendix: Agent Interaction

Figure 7 demonstrates the interaction between agents. Interactions within *Exanna* involve two agents, with the actor initiating actions based on its beliefs and goals and the observer responding to these actions. Beliefs are an agent's understanding of the world, which includes facts, information, observations, and assumptions. Partial observability can lead to agents developing distinct or even contradictory beliefs. Before the observer imposes a sanction, the actor presents rationales for its behaviors, which the observer then assesses.

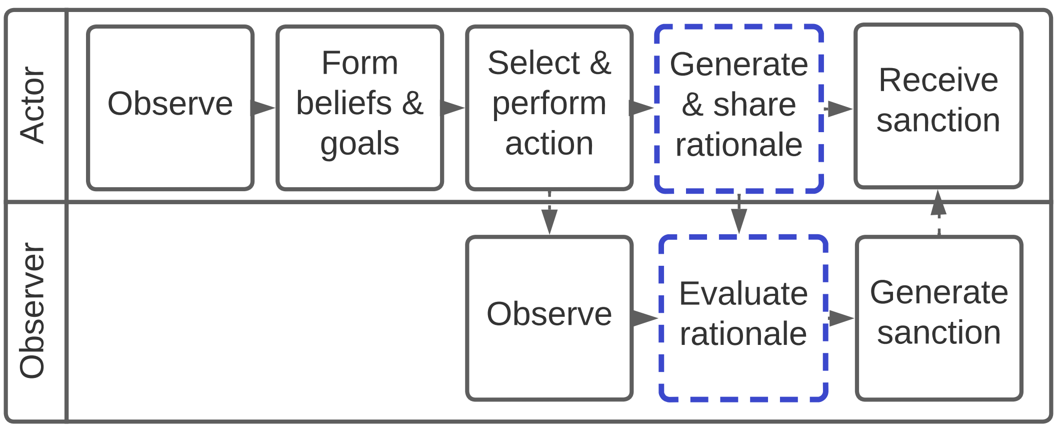


Figure 7: Interactions between *Exanna* agents. An agent forms its beliefs based on its observations of public contextual attributes and the beliefs revealed by other. The agent formulates goals according to its beliefs, subsequently utilizing these goals to make decisions and provide rationales for those decisions. Upon receiving the rationale, other agents evaluate and decide whether to accept it.

.2.2 Appendix: Procedures of XCS

The overall process of XCS includes the following sub-processes.

- Matching: A process that matches the current context and all rules/classifiers to generate a match set. For instance, in our running example , the match set for Bella may include (1) $\{\text{Risk} = \text{Low}\} \Rightarrow$ Wear [fitness = 0.3], (2) $\{\text{Risk} = \text{Low}\} \Rightarrow \neg$Wear [fitness = 0.7], (3)

$\{$OtherAgentType = Health$\}$ $\Rightarrow$ Wear [fitness = 0.8], and (4)$\{$OtherAgentType = Health$\}$ $\Rightarrow$ $\neg$Wear [fitness = 0.2]. The fitness is based on the accuracy of each rule's reward prediction.

- Covering: A process that guarantees diversity via adding a random classifier whose conditions match the current context. For instance, adding $\{$Risk = Low, Relationship = Friend$\}$ $\Rightarrow$ $\neg$Wear to the rule set.

- Action selection: XCS selects actions with pure exploration or pure exploitation with ε greedy. If not in exploration mode, this process returns the action with the highest fitness-weighted aggregation of reward.

$$\text{fitness}_a = \sum_{i}^{\text{rule}} \text{fitness}_i \times \text{numerosity}_i \times \text{predicted\_reward}_i \tag{1}$$

where a $\in A$ and A is the action space. Rules represent all rules applied to the context and for action a. With the above example and formula, the agent would choose not to wear a mask due to $\text{fitness}_{\neg\text{Wear}} > \text{fitness}_{\text{wear}}$.

- Formation of action set: It includes all classifiers that propose the chosen action based on the match set. For instance, $\{$Risk = Low$\}$ $\Rightarrow$ $\neg$ Wear, $\{$OtherAgentType = Health$\}$ $\Rightarrow$ $\neg$Wear, and $\{$Risk = Low, Relationship = Friend$\}$ $\Rightarrow$ $\neg$Wear.

- Updating classifier parameters (Urbanowicz and Browne 2017): An agent updates the rule parameters (e.g., accuracy and fitness) based on the received payoff. The following equation updates the predicted reward, where p is the predicted reward, β is the learning rate, and r is the received reward.

$$p \leftarrow p + \beta(r - p) \tag{2}$$

The prediction error ε is updated with the following equation.

$$\varepsilon \leftarrow \varepsilon + \beta(|r - p| - \varepsilon) \tag{3}$$

The fitness of a rule is based on its accuracy, which is inversely proportional to the prediction error. We update the accuracy $kappa$ with the following formula.

$$\kappa = \begin{cases} 1 & \text{if } \varepsilon < \varepsilon_0 \\ \alpha(\frac{\varepsilon}{\varepsilon_0})^{-\nu} & \text{otherwise,} \end{cases} \tag{4}$$

where α is the scaling factor that raises a non-accurate rule to be close to an accurate rule. ε_0

is the threshold of prediction error below which the prediction error of a rule is assumed to be zero. ν defines how accuracy is related to prediction error and aims to help differentiate similar classifiers. For fitness calculation, we next calculate the relative accuracy κ' of each rule.

$$\kappa' = \frac{\kappa}{\sum_{cl \in [A]} \kappa_{cl}} \tag{5}$$

where $[A]$ represents the corresponding action set. Finally, the fitness update of a rule is as follows.

$$F \leftarrow F + \beta(\kappa' - F) \tag{6}$$

where F is the fitness of a rule.

- Subsumption: A process that replaces offspring rules with more general parent rules if it exists. Otherwise, save the offspring rules. Specifically, a more general rule yields a minor prediction error. For instance, if rule {Risk = Low} $\Rightarrow$ $\neg$Wear has less prediction error than rule {Risk = Low, Relationship = Friend} $\Rightarrow$ $\neg$Wear, the former rule would replace the later rule and increases the numerosity.

- Deletion: Each action set has the same maximum number of rules. XCS removes the low-fitness rules.

.2.3 Appendix: Payoff Calculation with Values

Whereas preferences define the tendency of an individual to make a subjective selection among alternatives, values define the important things to an individual. Although both values and preferences are context-specific, values may transcend contexts (Liscio et al. 2021).

Each agent stores values in a tuple where each value maintains a corresponding $M_{individual}$. Since agents do not make decisions with single values but with tradeoffs among multiple related values, we aggregate value preferences when constructing a payoff (Ajmeri et al. 2020). Below, f is the aggregated payoff with all corresponding values after selecting strategy Rx when the other player selects strategy Cy from $M_{individual}$.

$$f = \sum_{i}^{values} v_i \times r_{RxCy} \tag{7}$$

We model interactions as games with payoffs f from the aggregation of $M_{individual}$.

.2.4 Appendix: Detailed Results

Table 2 summarizes the complete results.

$\mathbf{H}_{\textbf{Social Experience}}$ Social experience includes actor payoff and observer payoff. Figure 8 plots the payoffs of the actors who select actions, explain their behaviors, and receive feedback from observers in Share All, Share Decision Rules, and *Exanna* agent societies. Figure 9 shows the payoff from the observer who reacts to the actor's behavior in Share All, Share Decision Rules, and *Exanna* agent societies. The freedom-loving agents within *Exanna* society encounter more adverse feedback than other societies initially. However, some of them quickly adapt and begin to divert from their original goals. As a result of the behavioral change made by freedom-loving agents, there has been an enhancement in the feedback received by health-freak agents.

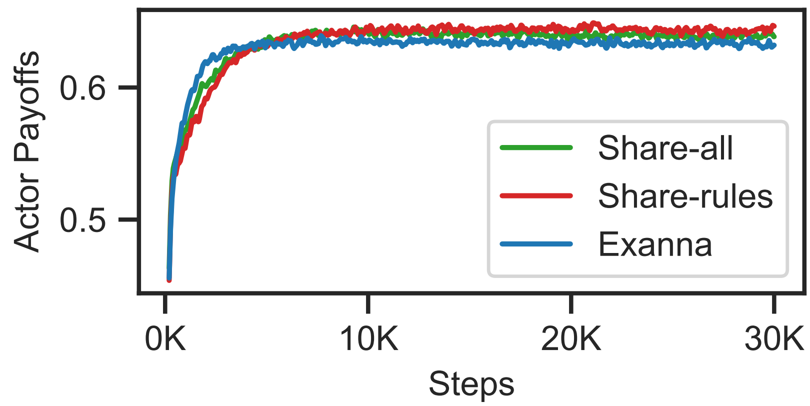
$\mathbf{H}_{\textbf{Flexibility}}$ We compare agents' flexibility of goals as the metric of evaluating $\mathrm{H}_{\mathrm{Flexibility}}$. Figure 10 compares $\mathrm{M}_{\mathrm{Flexibility}}$ by agent types in Share All, Share Decision Rules, and *Exanna* agent societies. Referring to Figure 9, the freedom-loving agents compromise on goals, thereby enhancing flexibility and enriching social experience.

.2.5 Appendix: Reproducibility Details

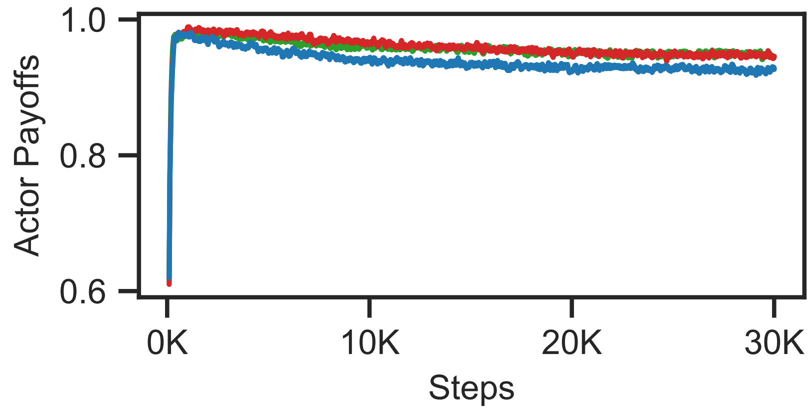
Table 3 lists the hyperparameters we set for our simulations, where we used the standard setting from (Urbanowicz and Browne 2017). The codebase of our simulation is available publicly.

.2.6 Appendix: Detailed Emerged Norms

Table 5 lists the norms that emerge in the simulations. An emerged norm is a rule adopted by more than 90% of agents in one society.



(a) Actor payoff for health-freak agents

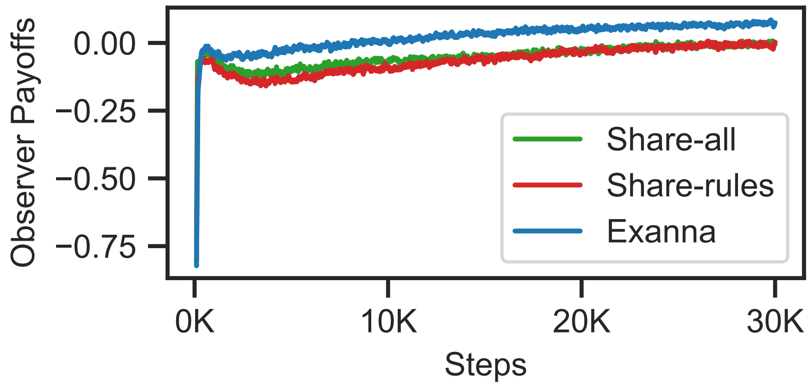


(b) Actor payoff for freedom-loving agents

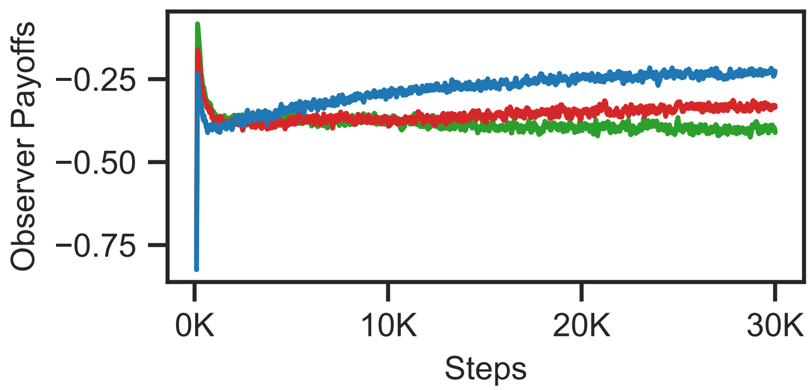
Figure 8: Comparing the actor payoff by agent types in various agent societies. Actors are agents who act and receive feedback from others. Health-freak agents in each society have similar actor payoffs. The freedom-loving agents in *Exanna* society have lower actor payoffs (Glass' $\Delta > 0.5$; $p < 0.001$) than the baseline societies.

Table 2: Results: Comparing mean ($\bar{X}$) and standard deviation (σ) social experience, resolution, privacy, and flexibility in various societies and agent types. p is p-value from t-test. M_{Social} has two subclasses, actor payoff and observer payoff.

		Share All	Share Rules	*Exanna*
$M_{Resolution}$	$\bar{X}$	0.58	0.58	0.60
	σ	0.02	0.01	0.02
	p	< 0.001	< 0.001	–
	Δ	1.80	3.11	–
M_{Social}	$\bar{X}$	0.59	0.62	0.70
	σ	0.06	0.02	0.05
	p	< 0.001	< 0.001	–
	Δ	1.80	3.11	–
Actor Payoff for healthy-freak	$\bar{X}$	0.63	0.63	0.63
	σ	0.03	0.03	0.03
	p	< 0.001	< 0.001	–
	Δ	10.40	9.62	–
Actor Payoff for freedom-loving	$\bar{X}$	0.96	0.96	0.94
	σ	0.03	0.03	0.03
	p	< 0.001	< 0.001	–
	Δ	0.60	0.64	–
Observer Payoff for healthy-freak	$\bar{X}$	-0.05	-0.07	0.02
	σ	0.09	0.09	0.09
	p	< 0.001	< 0.001	–
	Δ	2.71	2.39	–
Observer Payoff for freedom-loving	$\bar{X}$	-0.38	-0.36	-0.28
	σ	0.09	0.09	0.10
	p	< 0.001	< 0.001	–
	Δ	1.10	0.81	–
$M_{Privacy}$	$\bar{X}$	0.00	1.34×10^{-5}	0.25
	σ	0.00	2.11×10^{-5}	2.90×10^{-3}
	p	< 0.001	< 0.001	–
	Δ	∞	11 896.52	–
$M_{Flexibility}$	$\bar{X}$	0.10	0.09	0.12
	σ	0.01	0.01	0.03
	p	0.08	< 0.01	–
	Δ	1.59	2.89	–
Flexibility for healthy-freak	$\bar{X}$	0.14	0.12	0.13
	σ	0.06	0.07	0.06
	p	< 0.001	< 0.001	–
	Δ	0.85	0.44	–
Flexibility for freedom-loving	$\bar{X}$	0.06	0.06	0.09
	σ	0.03	0.03	0.03
	p	< 0.001	< 0.001	–
	Δ	0.90	0.87	–

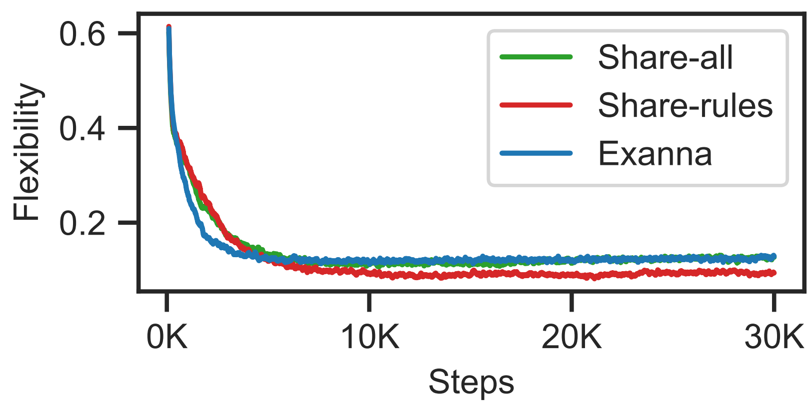


(a) Observer payoff for health-freak agents

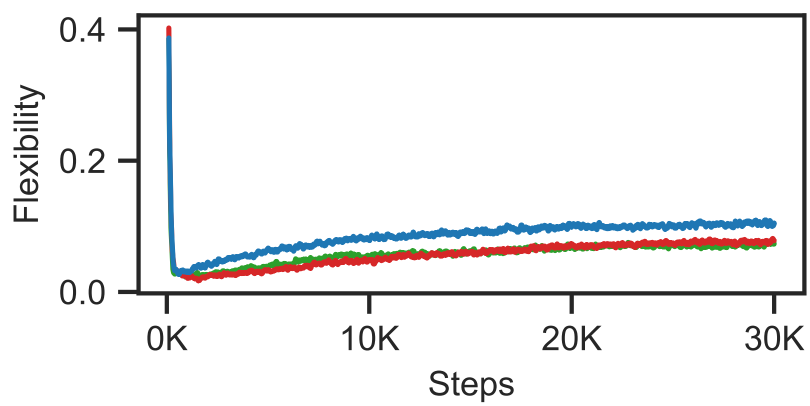


(b) Observer payoff for freedom-loving agents

Figure 9: Comparing the observer payoff by agent types in various societies. Observers give feedback based on observed behaviors and received rationales. The health-freak and freedom-loving agents in *Exanna* society have better observer payoffs (Glass' $\Delta > 0.8$; $p < 0.001$) than the baseline societies.



(a) Flexibility for health-freak agents



(b) Flexibility for freedom-loving agents

Figure 10: Comparing the flexibility by agent types in various agent societies. The freedom-loving agents in *Exanna* society has higher flexibility (Glass' $\Delta > 0.8$; $p < 0.001$) than the baseline societies.

Table 3: Hyperparameters for our settings.

Parameter	Value
Population size	200
Learning rate	0.1
Don't care probability	0.3
Accuracy threshold	0.01
Fitness exponent	5
Genetic algorithm threshold	25
Mutation probability	0.4
Crossover probability	0.8
Experience threshold for deletion	20
Experience threshold for subsumption	20
Fitness falloff	0.1

Table 4: Emerged norms in agent societies (1). Common means the norms emerge in each society.

Society	Norm			
	Premise			Consequence
Common	Risk preference OberverAgentType InteractWith location	= = = = =	NONE; ¬WEAR; HEALTH; COLLEAGUE; OFFICE	WEAR
	Risk preference OberverAgentType InteractWith location	= = = = =	NONE; ¬WEAR; HEALTH; COLLEAGUE; HOSPITAL	WEAR
	Risk preference OberverAgentType InteractWith location	= = = = =	RISK; ¬WEAR; HEALTH; COLLEAGUE; OFFICE	WEAR
	Risk preference OberverAgentType InteractWith location	= = = = =	RISK; ¬WEAR; HEALTH; COLLEAGUE; HOSPITAL	WEAR
Share All	Risk OberverAgentType InteractWith location	= = = =	NONE; HEALTH; COLLEAGUE; OFFICE	WEAR
Share Decision Rules	preference OberverAgentType InteractWith location	= = = =	¬WEAR; HEALTH; COLLEAGUE; OFFICE	WEAR

Table 5: Emerged norms in agent societies (2). Common means the norms emerge in each society.

Society	Norm: Premise	Consequence
Exanna	preference = ¬WEAR; InteractWith = COLLEAGUE; location = OFFICE	WEAR
	preference = ¬WEAR; InteractWith = COLLEAGUE; location = HOSPITAL	WEAR
	preference = ¬WEAR; ObserverAgentType = HEALTH; InteractWith = COLLEAGUE; location = OFFICE	WEAR
	preference = ¬WEAR; ObserverAgentType = HEALTH; InteractWith = COLLEAGUE; location = HOSPITAL	WEAR
	ObserverAgentType = HEALTH; InteractWith = COLLEAGUE; location = OFFICE	WEAR
	ObserverAgentType = HEALTH; InteractWith = COLLEAGUE; location = HOSPITAL	WEAR
	ObserverAgentType = FREEDOM; InteractWith = COLLEAGUE; location = HOSPITAL	WEAR
	Risk = RISK; ObserverAgentType = HEALTH; InteractWith = COLLEAGUE; location = OFFICE	WEAR
	Risk = NONE; ObserverAgentType = HEALTH; InteractWith = COLLEAGUE; location = OFFICE	WEAR